

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 06:01:15 ; Search time 5504.92 Seconds  
(without alignments)  
8515.190 Million cell updates/sec

Title: US-09-515-369b-1\_COPY\_1\_2240

Perfect score: 2240  
Sequence: 1 taatacgaactactatagg...aatctccacaagccttc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sta:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match Length	ID	Description
------------	-------	-------	--------------	----	-------------

1	2209	98.6	2256	9	AF217405	AF217405 Homo sapi	
2	2162.6	96.5	193317	2	AC098935	AC098935 Homo sapi	
c	3	2162.6	96.5	229302	2	AC023534	AC023534 Homo sapi
4	2105.4	94.0	180048	2	AC068122	AC068122 Homo sapi	
5	1826.2	81.5	11114	9	AY062931	AY062931 Homo sapi	
6	1319.2	58.9	7025	9	AF276916	AF276916 Homo sapi	
7	91.8	4.1	29034	9	HS243947	AJ243947 Homo sapi	
c	8	91.8	4.1	174210	2	AC108684	AC108684 Homo sapi
c	9	91.2	4.1	137769	9	AC002090	AC002090 Genomic S
10	90.8	4.1	193126	9	AL139275	AL139275 Human DNA	
11	90	4.0	177957	2	AC025298	AC025298 Homo sapi	
c	12	90	4.0	178477	2	AC022801	AC022801 Homo sapi
13	90	4.0	192695	2	AC098645	AC098645 Papio cyn	
14	89.2	4.0	181009	2	AC011060	AC011060 Homo sapi	
15	89.2	4.0	183483	9	AL161628	AL161628 Human DNA	
16	89.2	4.0	194590	2	HS15114	AL445623 Homo sapi	
17	88.4	3.9	168486	9	CNS057BN	AL163872 Human chr	
18	88.2	3.9	149973	2	AC023992	AC023992 Homo sapi	
19	88.2	3.9	158905	9	AC005666	AC005666 Homo sapi	
c	20	88.2	3.9	163242	2	AC068438	AC068438 Homo sapi
c	21	87.4	3.9	181134	2	AC074343	AC074343 Homo sapi
22	87	3.9	62497	2	AC103865	AC103865 Homo sapi	
c	23	87	3.9	178756	2	AC104420	AC104420 Homo sapi
c	24	86	3.8	116516	2	AC109809	AC109809 Homo sapi
c	25	85.4	3.8	86025	9	AL158175	AL158175 Human DNA
c	26	85.4	3.8	119361	2	AC092771	AC092771 Homo sapi
c	27	85.4	3.8	142192	2	AC092713	AC092713 Homo sapi
28	85.4	3.8	158341	2	AC021733	AC021733 Homo sapi	
29	85.4	3.8	160671	9	AC090451	AC090451 Homo sapi	
c	30	85.4	3.8	162625	9	AC006083	AC006083 Homo sapi
c	31	85	3.8	127936	2	AC022099	AC022099 Homo sapi
c	32	85	3.8	143536	2	AC020846	AC020846 Mus muscu
c	33	85	3.8	172830	2	AL137072	AL137072 Human DNA
c	34	84.6	3.8	59319	2	AC103842	AC103842 Homo sapi
c	35	84.6	3.8	114467	9	AP000462	AP000462 Homo sapi
36	84.6	3.8	154236	2	AC051650	AC051650 Homo sapi	
37	84.6	3.8	157322	9	AC012103	AC012103 Homo sapi	
c	38	84.6	3.8	157399	2	AP003167	AP003167 Homo sapi
c	39	84.6	3.8	170965	9	AB017654	AB017654 Homo sapi
c	40	84.4	3.8	144714	9	AC020891	AC020891 Homo sapi
41	84.4	3.8	145886	6	AP003438	AP003438 Homo sapi	
42	84.4	3.8	175830	2	AC090547	AC090547 Homo sapi	
c	43	84.4	3.8	188764	2	AC021238	AC021238 Homo sapi
c	44	84	3.8	170893	2	AC027537	AC027537 Homo sapi
c	45	83.8	3.7	37321	9	AC005498	AC005498 Homo sapi

ALIGNMENTS

RESULT	1	2256 bp	DNA	linear	PRI 19-SEP-2000
AF217405	AF217405	2256 bp	DNA	linear	PRI 19-SEP-2000
LOCUS	AF217405	2256 bp	DNA	linear	PRI 19-SEP-2000
DEFINITION	Homo sapiens suppression of tumorigenicity 16 protein (ST16) gene,				
PROMOTER	Promoter and partial sequence.				
ACCESSION	AF217405.1	GI:10185703			
VERSION	AF217405.1	GI:10185703			
KEYWORDS					
SOURCE					
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2256)				
AUTHORS	Madireddi, M.T., Dent, P. and Fisher, P.B.				
TITLE	AP-1 and c/EBP transcription factors contribute to mda-7 gene				
JOURNAL	promoter activity during human melanoma differentiation				
	J. Cell. Physiol. 185 (1), 36-46 (2000)				
PUBMED	20400556				
MEDLINE	10942517				
REFERENCE	2 (bases 1 to 2256)				
AUTHORS	Madireddi, M.T. and Fisher, P.B.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-DEC-1999) Urology, Columbia University, 630 West,				

168th Street, BB 1505, New York, NY 10032, USA

FEATURES  
source  
1..2256  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tisue\_type="Placenta"  
promoter  
1..2208  
/gene="ST16"  
1..>2256  
/gene="ST16"  
/note="mda-7"  
2209..>2256  
/gene="ST16"  
/product="suppression of tumorigenicity 16 protein"  
BASE COUNT 539 a 584 c 503 g 630 t  
ORIGIN

Query Match 98.6%; Score 2209; DB 9; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 tcaaccccttggaaaccagagctgcctgcctcccaaaagcttgtaactataactatcttca 91  
DB 1 TCACCTTTTGAACCCAGGCTCTGCTGCTCCCAAAAGCTGTACTCATTAAGATTCTCAA 60

QY 92 ctgaatctggccaaaggttcctagctgtctccctcctgaactcctcctcctgaagtaataatgc 151  
DB 61 CTGATTTTGGGCCAAGGTTCTCTAGGTTCTCTCTGACCTCTTGTGAAGTAATTAATGC 120

QY 152 tatgataagctcaatcggaagctgaaagccaaagcaaatgtttgcctgaactatccatgttta 211  
DB 121 TATGATTAAGCTCATCCGAGGCTGAGGCCCAAGGCACATGTTGGCTGCAATATCCATGTTA 180

QY 212 tatgattcctctcctcagaagatgagctactacagatcccaaggtgttacccctgaagccag 271  
DB 181 TATGATTTCTTCTCTCAGACAGAGTGAAGTACTACATCCAGATCCAGGTTTACCTTGAGGCCAG 240

QY 272 ccaagggtgatatacctgaactcagctcctcgtcccaagcctgcgccttaagaagctcatccca 331  
DB 241 CCAAGGTGATATCCATGACGTCATGCTGTTTCCAGCTGACCTGTTTAAACAGCTCATGCCA 300

QY 332 cctgcctgcctcctcccgctatactcgaaagaaagtaagcttgagattcaagctgccttgagg 391  
DB 301 CCTGCTGCTGCTCCCGCTATCTGCAAGACAGTAAGTCTAGGATTTTACGCTGCTGGGGG 360

QY 392 ctaatttccctcctcagctcctcgtcttagctgtctccgcgcctcccaactcactatct 451  
DB 361 CTGATTTTCCCTCTCAGCTTCCTGCTTTAGCTGTCTCCTGCCCTCCACATCACCATTACT 420

QY 452 ccagcaactcaccctgctctctctctcgtctcactacactgctccttgacaatcttatactc 511  
DB 421 CCAGCACTCTCAGCTGCTGTTCTTCTGCTCATCATGCTGCTGCTGACATCTTTTATCTC 480

QY 512 atagtgattagttgaaggtctctgttaatgcccataacacagtggttggaagggagg 571  
DB 481 ATGAGTGAATGTTAGGGGTTCTTGTAATGCCCTTAATCCACATGCTGGAAGGGGGAG 540

QY 572 ttggggaaagagatgagctctgagggtctgacctcctcctgaaggtgaagactcggagccct 631  
DB 541 TGGGGCAAGAGATGGGCTGTGGGCTGTGCCCTACTTCTGAGAGGTAAGACTGGGCCCT 600

QY 632 ccaggaacaaaggaattcaagctctgctgcaagctataagcaagcagactgctgccaaggaat 691  
DB 601 CCAAGGAACAAAGGATTACAGGCTGTGGGCAAGCTATAGCCAAAGACGACTGTGGCCAGGAT 660

QY 692 tgcagaaggaatcttctgtctctaaagaaataaacacacactgagtaataagagaggga 751  
DB 661 TGCAGAAAGGATATTTGTTGCTTAAAGAAATTAACAACCTGATGATGAGATGAGGAGGA 720

QY 752 ggggggtctctgctcagaaggaatctgggaagagctgcacaaaggtgtgcttactactcct 811  
DB 721 GGGGGGTGTTGGTCCAGAGAGATTTGGGAAGAGTCTGCCAAAGGGGTGTCTTACTACTCTCT 780

QY 812 cctcttcttctcaatccactgagctggaagcaagtaatcctgtcccccacgctacatc 871  
DB 781 CCTCTTTTCTTTCATCTCCACTGAGCTGAGGCAAGTTATCTGTGCCCAAGCTCATCTTC 840

QY 872 ctactccggttcccatagcctcggaccacaggttgggcaaacctctcctgtaagaagaccaga 931  
DB 841 CTACTCCGGTTTCCATCCATCCGAGCCAGGTTGGGCAAAACCTTCCGTAAAGCAACCGA 900

QY 932 caaggaactttttgagctctgtgccaatagtgctcaagctcaactcaactcactcctgcgc 991  
DB 901 CAGGAACATTTTATGAGCTGTGTGTGCTGATGCTGTGCTGACAACTACTATCTGTGCTG 960

QY 992 tctgtagcagaaagcaatlagcaacaalatgtcaacaacatatgtgaccccaatgaanaa 1051  
DB 961 TCTGTAGCACGAAGCAATTTAGCAACATATGTCAACAAACATATGATGACCCCATGAAAA 1020

QY 1052 cttaattatataatgatacggaaacacctgaaataatgtcttctcttggatttttcccca 1111  
DB 1021 CTATTATTATATATGATATCGGAACCTGAAATATATGCTTTTCTTGTGATTTTTCCTCA 1080

QY 1112 atcattaaaaaagcfaaaactactccttaggtcgcaaggttaagcaatctcactgaagc 1171  
DB 1081 ATCATTTAAAAACGTAAAAACCTTAGCTTAGTCCGAAGGTTTAAGCCATTTCTCAGCTTAGC 1140

QY 1172 agtggcaggctggatctggtctgtgacacagttgccaatccctcgaatcccaaatgtc 1231  
DB 1141 AGTGCAAGCTGATTTTGGCTTGTGACCTACAGTTGGGCAATTCCTGATTCCTCAAAATGT 1200

QY 1232 attcctcaggaagatgtggaacaaatacttatgggaagttgttgattaaagaagttaaagaag 1291  
DB 1201 ATTCTCTAGGGATGTGGCAATATCTTATGGAAGTCTGTGATTTAAACAGCTTAAGAG 1260

QY 1292 catcagaacatttccagaagcaggctagcacaatgcaggcgtctccttaactgaactcatgtga 1351  
DB 1261 CATCAGACATTTTCCAGGAGGGGCTGACATGCCAGGGCTCTCTAATGACCTCATTTGGA 1320

QY 1352 ttcactcgttttcaatggaagatcttgcaagaacaaatctcccaaacctaaagctctgagga 1411  
DB 1321 TTTCATCTGTTTTCATGAGAGATCTTTCGAAGACAAAGAAATTCCTCAACCTTGAAGTCTGAGA 1380

QY 1412 ctgtgctttgggaacacactgctcctctgtgataagcctcactgagccatgagtagaactctag 1471  
DB 1381 CTGTGCTTTGGGAAACATGCTCTCTGTGATGCCCTCATCTGGGCATATGTAAGTAATCAG 1440

QY 1472 agctgagtgctctgctagctgagaaataggtcagaagcctctgactgctgcctggagcgtctg 1531  
DB 1441 AGCTGAGTGCTTGTGCTAGCTGAGATAGGGTACAGAGCTCTTGACTGCCCTGGCAGTCTTG 1500

QY 1532 aacaataacgctgtctgtgtccctgagtggttcaagacccaagaagggcaagactaagccc 1591  
DB 1501 ACACATCAAGGCTGTGTGTGCTCCCTGAGTGTCTCAGAGCCACACAGGCCAAAGACTAGCCC 1560

QY 1592 accaagacccaagcctcccaactcttcgggcttctccatcgatcaatcttctatctt 1651  
DB 1561 ACCAGAGCACAGGCTCCCAAGCTTTCTGGGCTTGTCCATATGATGATATTTCTTATCTT 1620

QY 1652 cctggtttccagaacctaaggaaggaacatttggttgaatgataataacccataggaac 1711  
DB 1621 CCTGGTTTCCAGAACCTTAAGAGAGGACATTTTGTGATGATTAATTAACCTTGGAGAC 1680

QY 1712 catggttagctgcatgctcaagaaacacccccaactctccctggccctgtagtaataaagga 1771  
DB 1681 CATGGGTACTCTCATGTCAAGAAACATCTCCACCTTCTGCCCCCTGATGATTAATAAGGA 1740

QY 1772 gagggtactatcaaggttatttctctgctgtgaaactctgtcccaagcaatgaaatagggaccca 1831  
DB 1741 GAGGATCTTACAGCTTATTTCTTCTGCTGTGATCTGTCTCCAGCATTAATAGGCACTCA 1800

QY 1832 ttaattgaatlatcttgacaaggaagagaactggtgatactgtgcacagtaataatglatatta 1891  
DB 1801 TTATTGAATTAATTTTGACAGGAAGAGACTGCTGATGCTGACAGTAAATGATTAATTTA 1860











\* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces

	* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence as soon as it is available and the accession number W11791
	* be preserved.
	1 1400: contig of 1400 bp in length
*	1401 1500: gap of unknown length
*	1501 3388: contig of 1888 bp in length
*	3389 3488: gap of unknown length
*	3489 4879: contig of 1391 bp in length
*	4880 4980 6665: gap of unknown length
*	4980 6665: contig of 1986 bp in length
*	6966 7065: gap of unknown length
*	7066 9366: contig of 2301 bp in length
*	9367 11957: contig of 2491 bp in length
*	11958 12057: gap of unknown length
*	12058 14208: contig of 2151 bp in length
*	14209 17022: contig of 2714 bp in length
*	17023 17122: gap of unknown length
*	17123 20061: contig of 2939 bp in length
*	20062 20161: gap of unknown length
*	20162 24025: contig of 3864 bp in length
*	24026 26128: gap of unknown length
*	26129 26719: contig of 2593 bp in length
*	26719 28936: gap of unknown length
*	28936 32903: contig of 2118 bp in length
*	32903 32937 32781: gap of unknown length
*	32781 32881: gap of unknown length
*	32882 38895: contig of 6004 bp in length
*	38896 44306: gap of unknown length
*	44307 44406: contig of 5311 bp in length
*	44407 51589: gap of unknown length
*	51589 51689: contig of 7183 bp in length
*	51690 58591: gap of unknown length
*	58592 58691: contig of 6902 bp in length
*	58692 63442: gap of unknown length
*	63443 63542: contig of 4751 bp in length
*	63543 70554: gap of unknown length
*	70555 70654: contig of 7012 bp in length
*	70655 77794: gap of unknown length
*	77795 77894: contig of 7140 bp in length
*	77895 85841: gap of unknown length
*	85842 85941: contig of 7947 bp in length
*	85942 93227: gap of unknown length
*	93228 93327: contig of 786 bp in length
*	93328 100871: gap of unknown length
*	100872 100971: contig of 7544 bp in length
*	100972 112141: gap of unknown length
*	112142 112241: contig of 1170 bp in length
*	112242 126358: gap of unknown length
*	126359 126458: contig of 14117 bp in length
*	126459 141767: gap of unknown length
*	141768 141867: contig of 15309 bp in length
*	141868 155927: gap of unknown length
*	155928 156027: contig of 14060 bp in length
*	156028 180048: gap of unknown length
*	180049 24021: contig of 24021 bp in length.
FEATURES	
SOURCE	Location/Qualifiers
	1..180048 Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
BASE COUNT	49152 a 39956 c 40116 g 48114 t 2710 others
ORIGIN	
Query Match	94.0%; Score 2105.4; DB 2; Length 180048
Best Local Similarity	98.3%; Pred. No. 0;

Matches 2198; Conservative 0; Mismatches 16; Indels 23; Gaps 6;

QY 25 gaccgagtcacccctttgaaacccaggtcgcgcgtccaaagctgtacataactaga 84  
Db 80705 GGGTCGATACACCTTTGAAACCCAGGTCCTGCTCCAAACCTGTACTATACTAGA 80764  
QY 85 tctcaacatgagcttgaggccaaaggttcacaggttctctccctgaccccttcgaagta 144  
Db 80765 TTCTCACTAGATGTTGGCCCAAGTTCAGGTTCTCTCTGACCTTCCTTGAAAGTA 80824  
QY 145 ataatgcatatgataagctaaacgagagctgagagcccaagcaatatttgcttgaaatc 204  
Db 80825 ATATGCTATGATAGTACATGAGAGCTGAGGCCAGGACATGTTGGCTGAACATAC 80884  
QY 205 caagtataatgactcctcctccacagagatgagctacacagatcccaagtgtaacctg 264  
Db 80885 CATGTATATGATTCCTCTCCAGACAGATGAGTACTACATGATCCAGGTGATACCTG 80944  
QY 265 agggcagccaaagtgatatacagactcaagctcgtctccagcctgaccttaaacgt 324  
Db 80945 AGGCGACCCAAAGGTATGATGATGACTCATGCTCTGTTCCAGCTGCTTTAAACAGCT 81004  
QY 325 catccacacctgacctgacctcccg-ccatactgacagagatagatgaagattacagctgc 383  
Db 81005 CATCCCACTGCTGCTCCCTCCGACCTATGACAGACAGTATGCTAGATTTGAC-CCC 81063  
QY 384 cctggagagctcatcttcctccctcaggttccctcttaagctgtcctcctccacacac 443  
Db 81064 GCTGGGGCTCATTTTCCCTCCTCAGCTCTGCTTACGCTGCTCTCTCCCTCCACCTAC 81123  
QY 444 ctatctacacagactcaca-ccgtgctctctctctctgctacacagctcctctgacat 502  
Db 81124 CTATTACTCCAGACACTCTCATCTGCTCTCTCTCTGCTCATGACCTCTGTGACAT 81183  
QY 503 cctcatctcaatagatagatagaggtctctggtcaatgaccttaacatcagatgagga 562  
Db 81184 CTTATCTCATAGTACTAGTATGAGGCTCTTGTAAGCCCTTAATCAATGCTGGA 81243  
QY 563 aggggagagctgggggaaagagatgagctgaggctgtgacctctctggaaggtaaac 622  
Db 81244 AGGGGGAGTGGGGGAAAGAGAGCTGTGGGCTGCTCCACTCTCTGAGGGGTAAAGAC 81303  
QY 623 tcgggacctccaggaacaaagaa-ctcaggtctggtgacagctatagccaaagagactgtg 682  
Db 81304 TCGGGCCCTCCAGGAACAAAGATTCAGGCTGCTGACACTTACCCAAAGACACTGCTG 81363  
QY 683 gccagggatctgcaaaagagatcttctgcttaagaaataaacaacatgagatagag 742  
Db 81364 GCCAGGATTTGCAAGAGATATTTGTTGCTTAAAGAAATTAACACACTGAGTATGAG 81423  
QY 743 atggagagagaggggtctggtgagagagatctggaaagagctgcgaaggtgtcttct 802  
Db 81424 ATGGAGGAGGGGGGTGTGTGTCGCAAGAGATTTGGGAAGAGCTGCCAAAGGCTGTCTT 81483  
QY 803 actcaactcctccttctcttcaactcagctgagctgaggaagatlatctgtcccccac 862  
Db 81484 ACTCACT 81543  
QY 863 gtcaactctcactcccggttcccatgctgagcccaagattgggaaacatcttctctaa 922  
Db 81544 GTCACT 81603  
QY 923 agaaacagagagaaactattagagctgtgtgcaataatggtctcaatcaactatc 982  
Db 81604 AGAACCCAGACAGAGACTATTTAGGCTGTGTGCAATGTGCTCACTCAACTACTCTC 81663  
QY 983 atctctcctctctgagcaaaagagcaatagaacaatattgcaaaacataatgtaacc 1042  
Db 81664 ATCT 81723  
QY 1043 ccaatgaacacttataatataatgataagaaactgaaataatgcttcttcttctgalt 1102  
Db 81724 CCAATGAACCTTATTTATGATGATGAGAACTGAAATATATGCTCTCTCTCTCTCTCT 81783

QY 1103 tttccccaatcatataaaaaacgtaaaactactctttagtctgcaaggtttaagccattct 1162  
Db 81784 TTTTCCCAATCATATAAAAAAGTAAAACTACTCTTAAAGTGCAGAGGTAAAGCCATCT 81843  
QY 1163 caagcttaagctggcagagctgagatttgctgtgacctcaaglttgccaaatccctgaltc 1222  
Db 81844 CAGCTGACCACTGACAGCTGAGTATTTGGCTTGTGACCTACATGAGCCAAATCCGATTC 81903  
QY 1223 ccaaaatgactctcctcagagatgtggcgaataatcttaaggaagagctggatataaacga 1282  
Db 81904 CCAAAATGATTTCTCTCAGAGATGTGGCAATATCTTATGGAATGCTGATTAACAGA 81963  
QY 1283 gttaaagaacacagacatttccaggaagcggctagacacatgcaaggtcctcaatgac 1342  
Db 81964 GTTAAGAAGCATCAGACATTTCCAGGACGGGGCTAGCACATGCGCAGGGCTCTTAATGAC 82023  
QY 1343 ctcaatt-aggatcactcgttctcagagagatcttgcaagacagaagattcccaacctag 1401  
Db 82024 CTCAATTAGATTCATCTGTTTCATGAGATGAGATCTTGCAACAGAAATTCCTCAACCTAG 82083  
QY 1402 agctgaagactgtgcttgggaaacactgtctgtgaltgacctcaactgagcaatg 1461  
Db 82084 AGCTGAGAGACTGTGCTTGGGAAACACTGCTCTCTGATGCTCCCTCACTGGGCACATGG 82143  
QY 1462 tagaatctagaagctgaatgacctgtcagctggagatagagcttaagagctcttgacctg 1521  
Db 82144 TAGAATCTTAGAGCTGAGATGCCCTTGTGATGAGATAGGATGAGACTCTTGACTGCTCT 82203  
QY 1522 ggcagctctgacacatcaacagctgtctgtgctccctgagtggtctcaagacacaaagcca 1581  
Db 82204 GGCACTCTTGGACACATCACTCTCTGTGTGCTCCCTGAGTGTGAGACACACAGGCCA 82263  
QY 1582 agactagaccacacagagacacagagccctccagctctctgggtcttgcatactgacatt 1641  
Db 82264 AGACTAGGCCACACAGACACAGCCCTCCAGCTTCTGTGCTTGTGCTGCTGCTGCTGCT 82323  
QY 1642 ccttaactcctcctgtgttccagaaacctaaagagagacatttggtctgagtgattataa 1701  
Db 82324 CCTTATTTCTCTGTTTCCAGAACTTAAGAGAGGACATTTGTTGTTGATGATTAATA 82383  
QY 1702 cccctagagacataggtgagctgtcagctgcaagaaacactcctcaactcctgacctgtg 1761  
Db 82384 CCTTAGGACCATGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 82443  
QY 1762 gattaaagagaggtactcaaggttattcttctgctgtgactactgttcccaatgaa 1821  
Db 82444 GATTAAAGAGAGGTACTTACAGGTTATTTCTGCTGTGACTACTGT-CCAGCATGAA 82502  
QY 1822 taggcatcatattgattatttctgacaggaagagactgtgtgatactgacagtaat 1881  
Db 82503 TAGGSCATCATTTATTTATTTTGCACAGAAAGAGACTGTGATGCTGACAGATAT 82562  
QY 1882 aatgtatttacaatgtgtacagagcttcaacaagcctcgtgtgtgttctgtgtgtt 1941  
Db 82563 AATGATTTACTGTGTACAGAGTTTACCAAGACCTCTGCTGTTTCTCTCTCTCTCT 82622  
QY 1942 attaacctt-----gggacaatatlttaaatlttataatcaagaa 1983  
Db 82623 ATTACACTTGGCCCTTGGGGGAGGTGGACAAATTTTAAAAATTTTATATGACAGAGA 82682  
QY 1984 ctgcaagcagaagaagctaaagaaacttgccctgcccacacacagcagctgtagaagcctga 2043  
Db 82683 CTGCAAGCCGCAAGAGTAGAGACACTTGGCCCTGCCACACAGCCAGGTGATGAGCCGTA 82742  
QY 2044 actcaaaacccaggtctatctacaccccaagaggtctgttcccaactgctgattgtcctt 2103  
Db 82743 ACTCAAAACCCAGGTCTATCTCACTCAAGGCTGCTTTTCCCACTGCTGATGTTGCTTT 82802  
QY 2104 aaagtgtatgtgactagagcaatgaatlaattctctaaagaaagcattgacaaattccctt 2163  
Db 82803 AAAGTATGAGGTGCTAGCAATGAGATTAATCTCTAGAGAAAGCATATTCCTTCCCTT 82862

```
OY      2164  tctccaccctccctcttctccaccctcccccataagcccatatatatgcccacaa 2223
          |||
Db      82663  TCTCCACCTCCCTCTTTCTCCACACCCCTCCCGCATATATATATGCCCAA 82922
OY      2224  tctccacaagccttcg 2240
          |||
Db      82923  TCTCCACCAAGCCTTGC 82939

RESULT   5
LOCUS     AY062931                11114 bp    DNA       linear   PRI 26-NOV-2001
DEFINITION Homo sapiens Interleukin 24 (IL24) gene, complete cds.
ACCESSION AY062931
VERSION   AY062931.1  GI:17104816
KEYWORDS
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 11114)
           Rieder,M.J., Armet,T.Z., Carrington,D.P., Chung,M.-W., Lee,K.L.,
           Poel,C.L., Toth,E.J., Yi,Q. and Nickerson,D.A.
           Direct Submission
           Submitted (13-NOV-2001) Molecular Biotechnology, University of
           Washington, 1705 NE Pacific, Seattle, WA 98195, USA
           No cite this work please use: SeattleSNPs, NHLBI Program for
           Genomic Applications, UW-FHCRC, Seattle, WA (URL:
           http://pga.mbl.washington.edu).
FEATURES
     source
         1..11114
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /rpt_family="MERL_type"
     repeat_region
         543..856
         /rpt_type="dispersed"
         /rpt_type-dispersed
         627
         /frequency="0.01"
         /replace="a"
     variation
         634
         /frequency="0.10"
         /replace="c"
     variation
         806
         /frequency="0.05"
         /replace="g"
     repeat_region
         884..925
         /rpt_family="MERL_type"
         /rpt_type-dispersed
         1619..1620
         /frequency="0.07"
         /replace="a"
     repeat_region
         1625..1759
         /rpt_family="SINE"
         /rpt_type-dispersed
         1627
         /frequency="0.02"
         /replace="c"
     variation
         1889
         /frequency="0.06"
         /replace="g"
     mRNA
         join(1897..2068,2184..2329,3775..3970,4751..4813,5950..6108,6454..6528,7432..8320)
         /gene="IL24"
         /product="Interleukin 24"
         1897..8320
         /gene="IL24"
         2049
         /gene="IL24"
         /frequency="0.06"
         /replace="t"
     CDS
         join(2286..2329,3775..3970,4751..4813,5950..6108,6454..6528,7432..7515)
         /gene="IL24"

/codon_start=1
/product="Interleukin 24"
/protein_id="AAL34146.1"
/db_xref="GI:17104817"
/translation="MNFQORQDSMTLARPCPLLATASQNMVYLPCLGFTLLMS
VSGAGQGEHFHFGPCVGVVFOKLMEAFMAVKDTQKADQNTISARLDQEVQNSD
AESCYLVHRLTFEYKTVKFNHNTVEVRLKSFSTLANNFVLIVSQDPSQENMF
SIRDSAHRRPFLFRFAFRQLDVEALTFRLGEVDLILTMQKEYKI"
2324..2640
/rpt_family="SINE"
/rpt_type-dispersed
2461
/gene="IL24"
/frequency="0.90"
/replace="a"
2546
/gene="IL24"
/frequency="0.01"
/replace="a"
2625..3110
/gene="IL24"
/note="Region not scanned for variation"
3221
/gene="IL24"
/frequency="0.01"
/replace="c"
3260
/gene="IL24"
/frequency="0.37"
/replace="t"
3292
/gene="IL24"
/frequency="0.01"
/replace="a"
3352
/gene="IL24"
/frequency="0.01"
/replace="t"
3456
/gene="IL24"
/frequency="0.01"
/replace="a"
3517
/gene="IL24"
/frequency="0.02"
/replace="a"
3530
/gene="IL24"
/frequency="0.05"
/replace="a"
3551
/gene="IL24"
/frequency="0.99"
/replace="a"
3625
/gene="IL24"
/frequency="0.37"
/replace="c"
4061
/gene="IL24"
/frequency="0.01"
/replace="a"
4069
/gene="IL24"
/frequency="0.36"
/replace="t"
4082
/gene="IL24"
/frequency="0.02"
/replace="c"
4095..4167
/rpt_family="SINE"
/rpt_type-dispersed
```

	variation	4243-	/gene="IL24"	/frequency="0.37"	/replace="a"
	variation	4254	/gene="IL24"	/frequency="0.01"	/replace="g"
	variation	4332	/gene="IL24"	/frequency="0.64"	/replace="c"
	variation	4385	/gene="IL24"	/frequency="0.17"	/replace="a"
	variation	4602	/gene="IL24"	/frequency="0.38"	/replace="t"
	variation	4943	/gene="IL24"	/frequency="0.02"	/replace="a"
	variation	5020	/gene="IL24"	/frequency="0.01"	/replace="a"
	variation	5389	/gene="IL24"	/frequency="0.01"	/replace="a"
	variation	5392	/gene="IL24"	/frequency="0.01"	/replace="c"
	variation	5519	/gene="IL24"	/frequency="0.01"	/replace="c"
	repeat_region	5637..5681	/rpl_family="SINE"	/rpl_type="dispersed"	
	variation	6016	/gene="IL24"	/frequency="0.35"	/replace="c"
	variation	6020	/gene="IL24"	/frequency="0.02"	/replace="g"
	variation	6037	/gene="IL24"	/frequency="0.01"	/replace="c"
	repeat_region	6115..6280	/rpl_family="SINE"	/rpl_type="dispersed"	
	variation	6240	/gene="IL24"	/frequency="0.01"	/replace="c"
	variation	6282	/gene="IL24"	/frequency="0.37"	/replace="c"
	Query Match	81.5%	Score 1826.2;	DB 9;	Length 11114;
	Best Local Similarity	98.6%;	Pred. No. 0;	Mismatches	8; Indels 19; Gaps 2
	Matches 1870;	Conservative	0;	Mismatches	8; Indels 19; Gaps 2
Oy	363	gtatctcagagatttcgcgcacctggggcgctattttcccccctcaagcttctcgttagc	422		
Dd	1	GTACTCTAGGATTTCACGTCCCGTGGGGCCTCATTTTCCCTCAAGCTTCCTGCTTAGC	60		

OY	422	tgccccgagccctcccaactcaactcttaactccagacactctcaactgctctctctctc	482
OY	423	tgccccgagccctcccaactcaactcttaactccagacactctcaactgctctctctc	482
Db	61	tgctccgagccctcccaactcaactcttaactccagacactctcaactgctctctctc	120
OY	483	tcatactgagcctctbgaactcttaactccatagtagttagttaggggtctctgttaatgc	542
Db	121	tcatactgagcctctbgaactcttaactccatagtagttagttaggggtctctgttaatgc	180
OY	543	ccataatccacatgctgaggaagggaggtgagggaggaagatgcgcctgaggtctgc	602
Db	181	cctaaatccacatgctgaggaagggaggtgagggaggaagatgcgcctgaggtctgc	240
OY	603	ctactcttgagggtagagactcgggcccctccaggaacaaagatctcaggctgctgcgc	662
Db	241	ctactcttgagggtagagactcgggcccctccaggaacaaagatctcaggctgctgcgc	300
OY	663	latagcgaagcaagactctgctgcgaaggtattgcgaagaglatctgttgccttaagaa	722
Db	301	tatacgcgaagcaagactctgctgcgaaggtattgcgaagaglatctgttgccttaagaa	360
OY	723	taanaaacacgtgataagataggaaggaaggggttggtgcgaagaaatcttggaaga	782
Db	361	taanaaacacgtgataagataggaaggaaggggttggtgcgaagaaatcttggaaga	420
OY	783	gtctgcgaaggtgtgtcttaactaactctctctctctctcaactcaactgagctgag	842
Db	421	gtctgcgaaggtgtgtcttaactaactctctctctctctcaactcaactgagctgag	480
OY	843	gcagttatccgtctcccaagcgaactctcctctcctccgtcccaactgagccaggt	902
Db	481	gcagttatccgtctcccaagcgaactctcctctcctccgtcccaactgagccaggt	540
OY	903	tgaggaaactctctctctgtaagaacccaagaaactatcttaagctctgtgtccatat	962
Db	541	tgaggaaactctctctctgtaagaacccaagaaactatcttaagctctgtgtccatat	600
OY	963	gtctccagtcacaactctcaactctctgcccctctgacgaagaaagaacttagcaaatat	1022
Db	601	gtctccagtcacaactctcaactctctgcccctctgacgaagaaagaacttagcaaatat	660
OY	1023	gtcaaaaaactatgtgaccccaatgaaaacttatattatagttagaagaaactgaa	1082
Db	661	gtcaaaaaactatgtgaccccaatgaaaacttatattatagttagaagaaactgaa	720
OY	1083	ataabgtctctctctgttatcttctccccaactcaataaaaaogtaaaactctctcag	1142
Db	721	ataabgtctctctctgttatcttctccccaactcaataaaaaogtaaaactctctcag	780
OY	1143	tcgcaaaagttaagacatctcccaagcttagaagagtgagaggtctgtagttagcctac	1202
Db	781	tcgcaaaagttaagacatctcccaagcttagaagagtgagaggtctgtagttagcctac	840
OY	1203	agttggccaalccctgtaatcccaaaatgtaattccctcaagatgtgvgcaaatctatag	1262
Db	841	agttggccaalccctgtaatcccaaaatgtaattccctcaagatgtgvgcaaatctatag	900
OY	1263	gaagtgctctgataataaagaagtttaagaagatagaacatcttcagaagacggctagaat	1322
Db	901	gaatgtgcgtatttaaacagagtttaagaacatcagacatttcagacagcgctagacat	960
OY	1323	gcacagggtctctaaactgacactcaatt-gaatcaactgttctcaatgaggaactctgcaaga	1381
Db	961	gcacagggtctctaaactgacactcaatt-gaatcaactgttctcaatgaggaactctgcaaga	1020
OY	1382	caagaatctcccaaacctctagagctctgaggaactgtgctttggaacacactgctcgtctga	1441
Db	1021	caagaatctcccaaacctctagagctctgaggaactgtgctttggaacacactgctcgtctga	1080
OY	1442	tgccccaactgvggcacatggttagaatctagaactgaatgctgtctctagctvggaataaggg	1501
Db	1081	tgccccaactgvggcacatggttagaatctagaactgaatgctgtctctagctvggaataaggg	1140

QY 1502 tcagagcttctgagctgcccgcagctcttgcacacatcagctgtctgtccctcagtg 1561  
 DB 1141 TCAGAGCTCTTGAGCTGCCGCGCAGCTCTTGACATCAGCTGCTGTGTCCTGAGTG 1200  
 QY 1562 gttcagagccacacagagccaaagactagccacacagagccacagccctccagctcttg 1621  
 DB 1201 GTTCAGAGCCACACAGGCCAAGACTAGCCACGACGACACAGGCGCTCCAGCTTCTG 1260  
 QY 1622 gcttgcacatgcgtacatcttcccttctcctcctggttccagaacccaagaagagagcaca 1681  
 DB 1261 GCTGTGCTCAGCTGACATTTCTTATCTTCTTCTGTTCCAGAACTTAAAGAGAGGCGACA 1320  
 QY 1682 ttgtggttgcagtgatcttaaccctcagagagacatggtgagctgcatgctcagaaacactcc 1741  
 DB 1321 TTTTGTTGATGATGATTTTAACTTAAAGGACCATGAGTGGTGCATGTCAGAGAACTCC 1380  
 QY 1742 tcaacttccctgagccctgatgatgtaaaagagaggtacttcaaggtatcttccctgtg 1801  
 DB 1381 TCAACTTCCCTGGCCCTGATGATTAAGAGAGGAGTACTTACAGGTATTTCTTCTGCTGTG 1440  
 QY 1802 gactactgctccacacatgaagaagagcactatgtaataatttgagcagaagaagagact 1861  
 DB 1441 GACTACTGTCTCCACCATGAATAAGGACATTAATTAATTAATTTGACAGAGAGAGACT 1500  
 QY 1862 ggtgatgctgcacagtaataatgatttaccatgtgtacagagtttaccagacactctg 1921  
 DB 1501 GGTGATGCTGCACAGTAATAATGATTTACATGTGTACAGATTTTACCAAGACACTCTG 1560  
 QY 1922 tgtgttcttgccttctgttcttaccact-----ggagcaaatctt 1963  
 DB 1561 TGTGTTTGTGCTCTGTTTATTAACATTTGCCCTTGGGGAGGTGGGACAAATTTTAA 1620  
 QY 1964 aaattatgacatgagagagctgcagcagcagagaagcctaagaagcttccctgcacaca 2023  
 DB 1621 AAAATTTTAAATGACAGAGACTGACGCCAGACGAAAGCTAAAGAGACTTCCCTGCCACA 1680  
 QY 2024 cagccagtgtagagctgaactcaaacaccaggtctcaactcactcagaggtctgtctc 2083  
 DB 1681 CAGCCAGTGTAGAGCCCTGAACCTCAACCCAGGTCTCATCTCACCTCGGGCTGCTTTC 1740  
 QY 2084 cccatcgctgtatgtcttcaagaagtgagtgagctagcagaatgaatcttctctaga 2143  
 DB 1741 CCCATGCTGTATGTCTCTTAAAGTGAATGAGTGAAGCAATGAATTAATCTCTAGGA 1800  
 QY 2144 aagcagcacaattcccttctccacactccctcttcttcccaacccctcccatcag 2203  
 DB 1801 AAGCATGACCAATTTCCCTTCTCCACCTCCCTTCTTCCGCCACCCCTCCCATGAG 1860  
 QY 2204 ccccatatataatgcccaatctccacaagaagccttgc 2240  
 DB 1861 CCCCATATATATGCCCAATCTCCACAAAGCCCTTGC 1897

RESULT 6  
 AF276916 7025 bp DNA linear PRI 20-DEC-2000  
 LOCUS AF276916 Homo sapiens melanoma differentiation association protein 7 gene,  
 DEFINITION complete cds.  
 ACCESSION AF276916  
 VERSION AF276916.1 GI:11907920  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 7025)  
 AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.  
 TITLE The human MDA-7 gene  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 7025)  
 AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JUN-2000) Department of Surgery, University of

FEATURES  
 source Glasgow, Queen Elizabeth Building, Glasgow Royal Infirmary, Glasgow  
 G31 2ER, Scotland  
 Location/Qualifiers  
 1..7025  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 join(<1804..1847,3285..3480,4261..4323,5460..5618,  
 5964..6038,6942..7025)  
 /product="melanoma differentiation association protein 7"  
 join(1804..1847,3285..3480,4261..4323,5460..5618,  
 5964..6038,6942..7025)  
 /codon\_start=1  
 /product="melanoma differentiation association protein 7"  
 /protein\_id="AA041401.1"  
 /db\_xref="GI:11907921"  
 /translation="MNPORLQSLMTLARPPEPLAVASOMQVYLPCLGFTLLMS  
 QVSGAGQDFHFGPCVQKQKLEPMVAKDMQMDNTSARLLQDQVQVSD  
 ASCTIVHTLLEFLTKVFKNHNHTVEFKLSSTIANNFVLVSQLOPSQENEMF  
 STRDSNHRFLFRFAFQDVEALTLTALGEVDLLTMQRFKL"  
 BASE COUNT 1759 a 1668 c 1759 g 1838 t 1 others  
 ORIGIN

Query Match 58.9%; Score 1319.2; DB 9; Length 7025;  
 Best Local Similarity 97.5%; Pred. No. 0;  
 Matches 1379; Conservative 0; Mismatches 14; Indels 22; Gaps 3;  
 QY 848 taticgtcccccagctcacatccctcactccgcttcccaatgcctgagccaggttggc 907  
 DB 1 TATCTCTCCCCCAGCGACATTCCTACTCCCTTCCCATGCTGACCCAGCTTGGG 60  
 QY 908 aaactctctgtaaaagaccaggaactctttagctctgtgtccatagttct 967  
 DB 61 AAATCTTCTCTGTAAGAACCCAGACAGAACTTATTAAGCTGTGTCATATGCTCT 120  
 QY 968 cagtcacaacttactcatctctgtctctgttagcaggaagaacttagcaacaatgtcaa 1027  
 DB 121 CAGTCACACACTCATCTCTGCTCTGTAGCAGAAAGCAATTTAGCAATATGTCAA 180  
 QY 1028 caaacatagtgcaccccaagaaacttatttattatgtatgcgaaacctgaanaat 1087  
 DB 181 CAACATATGTGACCCCATGAACCTTATTTATTAAGTACGAAACCTGAATAAT 240  
 QY 1088 gtcttcttctgtatcttcccaatcaataaagaataaacttactttagtcgca 1147  
 DB 241 GTCTTCTTGTGATTTTCCCAATCATTTAAAGAGTAAACCTTCTTGTAGTGCA 300  
 QY 1148 aggttaagcaatctcagcttagcagtgagcagctgtgatttgcttctgtgacctacagt 1207  
 DB 301 AGGTTAAGCCATTTCTCAGCTTAGCAGTGGCAGCTGATTTGGCTGTGTGACCTAAGTTG 360  
 QY 1208 gccaatccctgattcccaaaatgtatctcctcaaggatgtgagcaaatcttattggaagt 1267  
 DB 361 GCCAATCCCTGATTCACCAAAATGTATTCCTCAGGAGTGTGGCAAAATATTATGGGCAAT 420  
 QY 1268 gctgtatlaaacagagtaagaagcagacatcttccagaagcgggttagcacatgcag 1327  
 DB 421 GCTGATTTAAACAGAGTTTAAAGACATTCAGACATTTCCAGAGCGGCTTACACATGCCAG 480  
 QY 1328 ggcctcctaacctgagcccaat-aggatcaltctgttcatgtgaagatcttgcagaaga 1386  
 DB 481 GGCTCTCTAATGACCTCATTTAGGATTCATCTGTTTCATGAGAGATCTTCAAGACAGA 540  
 QY 1387 attcctcaaacctagagctgtgagagctgtcttgggaaacactgtctcgttgatgcc 1446  
 DB 541 ATTCTCAAACTAGAGCTGAGAGCTGTGTTGGGAAACACGCTCTCTTGATGCC 600  
 QY 1447 tcaactggcacatgtagaagacatagagctgtgcttctgttagcttggagatagagtcaga 1506  
 DB 601 TCACTGGCACATGATGATAATCTAGAGCTGAGTCCCTTCTGACCTGAGATAGGCTCAGA 660  
 QY 1507 gctcttgccttgcctgagcagcttgcacacatcagcgtctgtgtccctcagtgatgtca 1566  
 |||||||

Db	661	GCCTTTGACATGCCCCCTGGCAGATCTTTGACACATCAACGCTGTCGTGTGTCCTCCCTGAAGTGATCA	720
QY	1567	gaagccacacagagccaaagactagaccaccagaagcagcagcctcccaagcttctcgtgcttg	1626
Db	721	GAGCCACACAGAGGCCCAAGACTAGCCCAACAGACAGCAGGCGCTCCAGCTTTCGGGCTTG	780
QY	1627	tgcagtcgtaacatttccctctatcttcttcctcgtgttcccgaaacctaaaggaagggacattttg	1686
Db	781	TCCATGCGTACATTTTCCCTTAATCTTCCTGGTTCCGAAGACCTAAGAGAGGACATATTTTG	840
QY	1687	gttcgtagtaataaaccctaaaggaacatggttagcgtcgtcgtatgltcagaanaacatccctcaac	1746
Db	841	GTTGAGATGATTTAATACCCCTAGGACCATGGGTAGCTGTCATGTGTAGGAAACATCCTCTCAAC	900
QY	1747	ttctctgagccctgtagtgaataaaggaagagtactacaggttacttcttcgcgtgt--aga	1803
Db	901	TTCTCTGGCCCTGTATGGATTTAAAGAAAGAGGTACTTACAGGTTAATTTCTTGCTATTACGA	960
QY	1804	ctacatctcccaagatlaataaggaagcaccatcatatgaatttatcttgacagaagaagaacttg	1863
Db	961	CTACGCTCCAGCATTAATAAGGSCATCATTTTGAATTTATTTTGACAGAGAGAGACTGG	1020
QY	1864	tgtatgctgcacagtaataatgtattctacatgltacagagtttaaccaagcactctgtg	1923
Db	1021	TGTATGCTCCACAGTAAATATGATTTATACATGTGTACAGAGTTACCAAGCAACCTCTGTG	1080
QY	1924	tgttttttgccttctgtttttttttaaacttggg-----acaatttttaa	1965
Db	1081	TGTGTTTTCCCTCTGTTTATTTACACTGTCGCCCTTGGGGGAGAGGTGGGACAAATTTTAA	1140
QY	1966	aattatatacagcagaagactcagcagcagagaagaactaaagaacttgcccttgccacacaca	2025
Db	1141	AATTTATACATCCAGAGACTGCGAGGCAAGAGAAAGTAAAGACTTGGCCCTGCCACACA	1200
QY	2026	gcacagtgagagagccctgaactcaaaacccagagctcatctcacctaaaggagcgtcttccc	2085
Db	1201	GCCAGTGGTAGAGCCTGAACCTCAAAACCGGCTTATCTCACCTCAAGGGGGCTGCTTTCC	1260
QY	2086	catcgctgatatgtcccttaagaatgaatgggtgactaggaatgaatgaatcttctctagaaa	2145
Db	1261	CATCGCTGAATTTGCCCTTAAAGTGAATGGGTGACTAGGCAATGAAGTAAATTCCTTAGAAA	1320
QY	2146	gaatgaaccaatttccctttctcacactccctcttcttctccacccctcccccatacgcc	2205
Db	1321	GCATGACCAATTTCCCTTTCTCCACCTCCCTCTTTTCTCTCCACCCCTCCCATCACGCC	1380
QY	2206	cccatatataatgccaaaatctccacaagaagccttgc	2240
Db	1381	CCCATATATATATGCCCAAAATCTCCACAAAGCGCTGC	1415
RESULT	7		
LOCUS	HSA243947	29034 bp	DNA
DEFINITION	HSA243947	29034 bp	linear
ACCESSION	AJ243947		
VERSION	AJ243947.1	GI:5912555	
KEYWORDS	VGX-A gene.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li,X.M., Yen,P.H. and Shapiro,L.J.		
JOURNAL	Characterization of a low copy repetitive element S332 involved in		
MEDLINE	the generation of frequent deletions of the distal short arm of the		
REFERENCE	human X chromosome		
AUTHORS	Nucleic Acids Res. 20 (5), 1117-1122 (1992)		
	2 (bases 1 to 29034)		
	Fukami,M., Kirsch,S., Schiller,S., Richter,A., Benes,V., Franco,B.,		
	Muroya,K., Rao,P., Merker,S., Niesler,B., Ballabio,A., Ansoorge,W.,		
	Ogata,T. and Rappold,G.A.		

TITLE	A member of a gene family on xp22.3, VCX-A, is deleted in patients with X-linked nonspecific mental retardation		
JOURNAL	Am. J. Hum. Genet. 67 (3), 563-573 (2000)		
MEDLINE	20395828		
REFERENCE	3 (bases 1 to 29034)		
AUTHORS	Richier, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-AUG-1999) Richier A., Biochemical Instrumentation, EMBL, Meyerhofstrasse 1, Heidelberg, 69117, GERMANY		
FEATURES	location/Qualifiers		
source	1. 29034		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="Xp22.3"		
	/map="between 40BT and DXS1139"		
	/clone="LNC110"		
	/sex="male"		
	/cell_line="lymphoblastoid"		
	/clone_lib="Lawrence Livermore X-chromosome specific library"		
	/note="Variably charged, X chromosome mRNA on CRI-S232A"		
exon	10574..10732		
	/gene="VCX-A"		
	/number=1		
gene	10574..11947		
	/gene="VCX-A"		
	10733..11046		
	/gene="VCX-A"		
Intron	/number=1		
	11047..11294		
	/gene="VCX-A"		
	/number=2		
CDS	join(11195..11294,11487..11947)		
	/gene="VCX-A"		
	/codon_start=1		
	/product="VCX-A protein"		
	/protein_id="CAB36197.1"		
	/db_xref="GI:5912556"		
	/translation="MSPKPRASGPAKATEAGKRKSSQSPSPDKKKTKVARKGKA VARGRRKGKATKMAATAPAEESGPAAPGSDPSQELPQHLELPPEEPVSEGTQHD PLPSQSELEEPLOSSEVEEPLSQSEVEEPLSQSEVEEPLSQSEVEEPLSQSEVE EPLSQSEVEEPLSQSEMEELPSV"		
Intron	11295..11486		
	/gene="VCX-A"		
	/number=2		
exon	11487..11947		
	/gene="VCX-A"		
	/number=3		
repeat_region	11697..11936		
	/citation=[1]		
	/rpt_family="RUI"		
repeat_region	12097..12297		
	/citation=[1]		
	/rpt_family="RUI"		
BASE COUNT	9123 a 5623 c 5586 g 8631 t 71 others		
ORIGIN			
Query Match	4.1%; Score 91.8; DB 9; Length 29034;		
Best Local Similarity	75.5%; Pred. No.14;		
Matches 114; Conservative	0; Mismatches 37; Indels 0; Gaps 0;		
QY	891 ctggagcccaaggttgaggcaactctccgttaagagaccagacaggaactatlttagctc 950		
DB	4201 CTAAATCAGAGTTCTGCACCTTTCTGTAAAGCAGCAGACGAACTATTTAGGCTT 4260		
QY	951 tgtgtgcacatagtctcagtcacaacacacalcactctgcctctgttagcagaaagcaat 1010		
DB	4261 TGAAGCAATATGCGCTCTTTTACAACTACATCTCGCTTGTAGCAAAAAAACAAGC 4320		
QY	1011 tagcaacaatatgtcacaacaacatagtgcac 1041		
DB	4321 CACAGACAATATGTTAAATTAATTAATGTGGC 4351		



RESULT 8  
AC108684/c  
LOCUS  
DEFINITION  
SEQUENCE: 4 unordered pieces.  
AC108684  
AC108684.1 GI:18449745  
HTG: HTGS\_PHASE1: HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Obiola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralungu,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barberia,J., Benton,U., Blmage,K., Blankenburg,K., Bonnin,D.,  
Bouch,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,  
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,J.F., Howard,S., Huber,J., Huliyil,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawlin,E., Meleod,M.P., Meleod,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,S., Ogih,M., Okunishi,S.,  
Nguyen,N., Nickerson,E., Nwokweto,S., Ogih,M., Okunishi,S.,  
Ogungbe,N., Oviedo,R., Pace,A., Payton,B., Peery,U., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Severy,G.,  
Scheer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinton,R., Wall,R.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,  
Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (31-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
-----  
Center: Baylor College of Medicine  
Genome Center  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
-----  
Project Information  
Center project name: HDOM  
Center clone name: RP11-359020  
-----  
Summary Statistics  
Sequencing vector: Plasmid; M77789

Chemistry: Dye-terminator Big Dye: 1008 of reads  
Assembly program: Phrap: version 0.990329first call to  
findPhrapList  
Consensus quality: 170951 bases at least Q40  
Consensus quality: 171318 bases at least Q20  
Consensus quality: 171608 bases at least Q20  
Estimated insert size: 170939; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 9.9x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drift\_data.html).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N. Gaps between the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 76682: contig of 76682 bp in length  
\* 76683 76782: gap of unknown length  
\* 76783 150676: contig of 73894 bp in length  
\* 150677 150776: gap of unknown length  
\* 150777 164461: contig of 13685 bp in length  
\* 164462 164561: gap of unknown length  
\* 164562 174210: contig of 9649 bp in length.  
Location/Qualifiers  
source  
1. 174210  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="Xp"  
/clone="RP11-359020"  
BASE COUNT 52809 a 34893 c 34663 g 51537 t 308 others  
ORIGIN  
Query Match 4.1%; Score 91.8; DB 2; Length 174210;  
Best Local Similarity 75.5%; Pred. No. 1.6e-14;  
Matches 114; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 891 ctggaccacaggtggcgaactcttcgttaagaacgaagaacatttagctc 950  
Db 159418 CTAAATCAGAGTCTCTCAACCTTTCCTGTAAGGACGACGAGGATTTTGA6GCTT 159359  
QY 951 tctgtccatcatgtctcagtcacactcaactcctctcctctgtagcagaagaacat 1010  
Db 159358 TGAAGCAATATGCTCTCTTTTACACTACTATCTCTCTCTTGTGACAAAAAAGACG 159299  
QY 1011 tagcaacaatatgtcaacaacatatgtgac 1041  
Db 159298 CACGACAAATATGTAATTAATTAATGATGTGAC 159268  
RESULT 9  
AC002090/c  
LOCUS  
DEFINITION  
AC002090  
AC002090.1 GI:2160130  
HTG.  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Hawkins,T.L., Birren,B.W., Fasnman,K.H., Nussbaum,C. and Lander,E.S.  
Genomic sequence from Human 17  
Unpublished  
2 (bases 1 to 137769)  
Hawkins,T.L., Reeve,M.P., Christofferson,A., Birren,B.W.,  
Fasnman,K.H., Lander,E.S., McKernan,K., Munro,C., Richardson,P.,  
Barna,N., Chang,A., Cooke,P., Daly,M.J., Forrest,C., Frapp,W.J.,

TITLE	Gage, D., Geraigery, K., Hagos, B., Jacotot, L., Lane, M., Mackenzie, J., Margulis, N., McDermott, J., Moloney, N., Morrow, J., Naylor, J., O'Connor, T., Peterson, K., Rollins, G., Spencer, J., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Wilmer, F., and Zody, M.
JOURNAL	Submitted (13-MAY-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 137769)
AUTHORS	Hawkins, T. L., Reeve, M. P., Christoffersen, A., Birren, B. W., Fasmann, K. H., Lander, E. S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M. J., Devon, K., Dewart, K., Forrest, C., Gage, D., Geraigery, K., Hagos, B., Harris, K., Huang, J., Hul, L., Jacotot, L., Kirby, A., Lane, M., Mackenzie, J., Margulis, N., McDermott, J., Molla, M., Morrow, J., Nachman, A., Naylor, J., Nusbaum, C., O'Connor, T., Olotu, A., Peterson, K., Reeve, M. P., Roberts, D., Rollins, G., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Wilmer, F., Zemtseva, I., and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (02-SEP-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	5 (bases 1 to 137769)
AUTHORS	Hawkins, T. L., Birren, B. W., Fasmann, K. H., Nusbaum, C., Lander, E. S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M. J., Devon, K., Dewart, K., Forrest, C., Gage, D., Geraigery, K., Hagos, B., Harris, K., Huang, J., Hul, L., Jacotot, L., Kirby, A., Lane, M., Mackenzie, J., Margulis, N., McDermott, J., Molla, M., Morrow, J., Nachman, A., Naylor, J., Nusbaum, C., O'Connor, T., Olotu, A., Peterson, K., Reeve, M. P., Roberts, D., Rollins, G., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Wilmer, F., Zemtseva, I., and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (09-SEP-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	6 (bases 1 to 137769)
AUTHORS	Hawkins, T. L., Birren, B. W., Fasmann, K. H., Nusbaum, C., Lander, E. S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M. J., Devon, K., Dewart, K., Forrest, C., Gage, D., Geraigery, K., Hagos, B., Harris, K., Huang, J., Hul, L., Jacotot, L., Kirby, A., Lane, M., Mackenzie, J., Margulis, N., McDermott, J., Molla, M., Morrow, J., Nachman, A., Naylor, J., Nusbaum, C., O'Connor, T., Olotu, A., Peterson, K., Reeve, M. P., Roberts, D., Rollins, G., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Wilmer, F., Zemtseva, I., and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (06-JUN-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jun 4, 1997 this sequence version replaced g1:2085774. The Staden databases, finishing information, and all chromatographic files used in the assembly of this clone are available from our anonymous ftp site.
FEATURES	All repeats were identified using RepeatMasker: Smlt, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html.
SOURCE	Location/Qualifiers 1. 137769 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /clone="4314"
repeat_region	/clone_1lib="Research Genetics/Cal Tech CITB9785K-B (plates 1-194)"
repeat_region	9...113
repeat_region	/rpt_family="MUTIC"
repeat_region	114..416
repeat_region	/rpt_family="Alusx"
repeat_region	418..660
repeat_region	/rpt_family="MLTIC"
repeat_region	951..1021
repeat_region	/rpt_family="(CAT)n"
repeat_region	complement(1659..1701)
repeat_region	/rpt_family="(CA)n"
repeat_region	2009..2282
repeat_region	/rpt_family="LINE2"
repeat_region	complement(2539..2700)
repeat_region	/rpt_family="MIR"
repeat_region	complement(2709..2884)
repeat_region	/rpt_family="LINE2"
repeat_region	complement(3877..4032)
repeat_region	/rpt_family="LINE2"
repeat_region	4630..4841
repeat_region	/rpt_family="MIR"
repeat_region	complement(5998..6079)
repeat_region	/rpt_family="(GA)n"
repeat_region	complement(7528..7829)
repeat_region	/rpt_family="Alusx"
repeat_region	complement(8692..8879)
repeat_region	/rpt_family="MIR"
repeat_region	9321..9667
repeat_region	/rpt_family="LINE2"
repeat_region	complement(9688..9982)
repeat_region	/rpt_family="Alusg"
repeat_region	9985..10210
repeat_region	/rpt_family="LINE2"
repeat_region	11052..11216
repeat_region	/rpt_family="MIR"
repeat_region	complement(12296..12393)
repeat_region	/rpt_family="(GGG)n"
repeat_region	12572..12644
repeat_region	/rpt_family="AT_rich"
repeat_region	12908..12995
repeat_region	/rpt_family="LINE2"
repeat_region	complement(14208..14363)
repeat_region	/rpt_family="LINE2"
repeat_region	complement(14577..14686)
repeat_region	/rpt_family="LINE2"
repeat_region	14837..14904
repeat_region	/rpt_family="(CAGC)n"
repeat_region	complement(14906..14963)
repeat_region	/rpt_family="LINE2"
repeat_region	complement(17782..17841)
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(19230..19258)
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(19446..19742)
repeat_region	/rpt_family="AluY"
repeat_region	19749..19785
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(20413..20518)
repeat_region	/rpt_family="MER5B"
repeat_region	21555..21649
repeat_region	/rpt_family="MER5A"
repeat_region	complement(21855..21892)
repeat_region	/rpt_family="LINE2"
repeat_region	complement(22871..22908)
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(23129..23168)
repeat_region	/rpt_family="LINE2"
repeat_region	complement(23296..23654)
repeat_region	/rpt_family="LINE2"
repeat_region	23894..24398
repeat_region	/rpt_family="MLT2CA"
repeat_region	complement(24536..24675)



misc_feature	complement(35079, .35571)
misc_feature	/note="match": GSS: Em:AQ692857"
misc_feature	/note="match": GSS: Em:AQ692857"
misc_feature	/note="match": GSS: Em:AQ892878"
misc_feature	.41966, .42582
misc_feature	/note="match": GSS: Em:AQ308209"
misc_feature	.41976, .42381
misc_feature	/note="match": GSS: Em:AQ124355"
misc_feature	.46366, .46984
misc_feature	/note="match": GSS: Em:AQ882156"
misc_feature	.47399, .47848
misc_feature	/note="match": GSS: Em:AQ811844"
misc_feature	.64782, .65002
misc_feature	/note="match": GSS: Em:AQ772609"
misc_feature	.64783, .65002
misc_feature	/note="match": GSS: Em:AQ683445"
misc_feature	/complement(64835, .64983)
misc_feature	/note="match": STS: Em:G08189"
misc_feature	/complement(64864, .64983)
misc_feature	/note="match": GSS: Em:AL328221"
misc_feature	/complement(64864, .64975)
misc_feature	/note="match": GSS: Em:AL324222"
misc_feature	/complement(64888, .64983)
misc_feature	/note="match": GSS: Em:AQ932064"
misc_feature	.64988, .65015
misc_feature	/note="match": STS: Em:G09742"
misc_feature	/complement(64898, .64990)
misc_feature	/note="match": GSS: Em:A2257084"
misc_feature	/complement(64898, .64899)
misc_feature	/note="match": GSS: Em:AQ674454"
misc_feature	/complement(64900, .64983)
misc_feature	/note="match": GSS: Em:A227938"
misc_feature	.64903, .65013
misc_feature	/note="match": GSS: Em:AL191898"
misc_feature	.64903, .64991
misc_feature	/note="match": GSS: Em:AL309316 Em:AL316981 Em:AL346021"
misc_feature	.64903, .64993
misc_feature	/note="match": GSS: Em:AL224697"
misc_feature	/complement(64903, .64991)
misc_feature	/note="match": GSS: Em:AL284370"
misc_feature	/complement(64903, .64983)
misc_feature	/note="match": STS: Em:G10109"
misc_feature	.64905, .65002
misc_feature	/note="match": GSS: Em:AQ284087"
misc_feature	/complement(74424, .74751)
misc_feature	/note="match": GSS: Em:AQ151384"
misc_feature	/complement(77154, .77547)
misc_feature	/note="match": GSS: Em:AQ192713"
misc_feature	.80797, .81303
misc_feature	/note="match": GSS: Em:AQ606690"
misc_feature	/complement(80886, .81417)
misc_feature	/note="match": GSS: Em:AQ602297"
misc_feature	/complement(91202, .91756)
misc_feature	/note="match": GSS: Em:AQ063130"
misc_feature	/complement(102120, .102591)
misc_feature	/note="match": GSS: Em:AQ176699"
misc_feature	.107123, .107721
misc_feature	/note="match": GSS: Em:AQ524997"
misc_feature	/complement(115935, .116552)
misc_feature	/note="match": GSS: Em:AQ080194"
misc_feature	/complement(116067, .116478)
misc_feature	/note="match": GSS: Em:AQ031004"
misc_feature	/complement(116195, .116555)
misc_feature	/note="match": GSS: Em:B32716"
misc_feature	/complement(128470, .128806)
misc_feature	/note="match": GSS: Em:B36077"
misc_feature	/complement(128811, .129176)
misc_feature	/note="match": STS: Em:H5164XF4"
misc_feature	.139541, .139825
misc_feature	/note="match": GSS: Em:AQ278260"
misc_feature	.163778, .165946
misc_feature	/note="match": STS: Em:G26303"
misc_feature	.171874, .172092

	/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."
misc_feature	/note="match: GSS: Em:AQ82642"
misc_feature	/note="match: GSS: Em:AQ374550"
misc_feature	/note="match: GSS: Em:AQ131194"
misc_feature	/note="match: GSS: Em:AQ148743"
misc_feature	/note="match: GSS: Em:AQ420001"
misc_feature	/note="match: GSS: Em:AQ545214"
BASE COUNT	54062 a 43323 c 43111 g 52630 t
ORIGIN	
Query Match	4.1% Score 90.8; DB 9; Length 193126;
Best Local Similarity	55.7%; Pred No. 3.1e-14;
Matches 194; Conservative	0; Mismatches 152; Indels 2; Gaps 1;
QY 901	gttgggcaactcttcctcgttaaagaaccaggaacattatcttaggcctcgtcgccat 960
Db 125379	GGTCACCAAACTTTTCTGTAAAGAACGATAGCATATTATTTAGCGTTTGCCGCAT 125438
QY 961	atggctcagtacaacctcatctctctcctcttgtagcagaagaacattagccaat 1020
Db 125439	ATAATCTCGTACACACGCTCACCCTCCGCCATTCGACAGTGAAMAGACTTAGACAAA 125498
QY 1021	atgtcaacaacatalgtgccccalgaanaacttatattatgatagcaagaacctga 1080
Db 125499	ACGTAATGACAGGGGCGATGCGTGCCCCACAGTAAACTTTATAGGGCACCATGAAATTTTAA 125558
QY 1081	aataatgctctcttttgtatctttcccccaatcaaaaaaacgtaaaactactctta 1140
Db 125559	TTTTGATGTGTACCAAAGTATTTTCTTTGATTTTCCAACCCAAATTAATAAATGTAAAAAC 125618
QY 1141	gttcgcaagggttaagccattctcagcttagcagtgacggcgctgaattggtcttgacct 1200
Db 125619	CTTCTCTTAAGTTGAGGATGAGAACAACAAACG--GCATCGGGGCTGGATTTGGCTCATGGGCC 125676
QY 1201	acagttggccaatcccgtattcccaaaaatgtatcttcagaagatctgg 1248
Db 125677	ATTGTGTGCCGACCCCTGATTGACATCATCCCATCTCACAATGTGG 125724
RESULT 11	
LOCUS	AC025298 177957 bp DNA linear HTG 03-JUN-2000
DEFINITION	Homo sapiens clone RP11-2N21, WORKING DRAFT SEQUENCE, 5 unordered pieces.
ACCESSION	AC025298
VERSION	AC025298.3 GI:825254
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Bitren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Unpublished
JOURNAL	2 (bases 1 to 177957)
REFERENCE	Anderson,S., Baldwin,J., Barra,N., Bastien,V., Beda,F., Boguslavsky,I., Boukhalter,B., Brown,A., Burkett,G., Campionio,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,W., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,







TITLE  
JOURNAL  
COMMENT

Myman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 12, 2000 this sequence version replaced g1:6984405.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L3156

Center clone name: L3156

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 16361 bases at least Q40

Consensus quality: 172073 bases at least Q30

Consensus quality: 175479 bases at least Q20

Insert size: 15500; agarose-ff

Quality coverage: 5.1 in Q20 bases; agarose-ff

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1504: contig of 1504 bp in length  
1505: gap of 100 bp  
1605: 2791: contig of 1187 bp in length  
2792: 2891: gap of 100 bp  
2892: 3909: contig of 1018 bp in length  
3910: 4009: gap of 100 bp  
4010: 5556: contig of 1547 bp in length  
5557: 5656: gap of 100 bp  
5657: 7482: contig of 1826 bp in length  
7483: 7582: gap of 100 bp  
7583: 8926: contig of 1344 bp in length  
8927: 9026: gap of 100 bp  
9027: 10470: contig of 1444 bp in length  
10471: 10570: gap of 100 bp  
10571: 12199: contig of 1629 bp in length  
12200: 12299: gap of 100 bp  
12300: 14087: contig of 1788 bp in length  
14088: 14187: gap of 100 bp  
14188: 15707: contig of 1520 bp in length  
15708: 15807: gap of 100 bp  
15808: 17945: contig of 2138 bp in length  
17946: 18045: gap of 100 bp  
18046: 20017: contig of 1972 bp in length  
20018: 20117: gap of 100 bp  
20118: 22115: contig of 1998 bp in length  
22116: 22215: gap of 100 bp  
22216: 24556: contig of 2341 bp in length  
24557: 24656: gap of 100 bp  
24657: 27568: contig of 2912 bp in length  
27569: 27668: gap of 100 bp  
27669: 32527: contig of 4859 bp in length  
32528: 32637: gap of 100 bp  
32638: 37892: contig of 5265 bp in length  
37893: 37992: gap of 100 bp  
37993: 41265: contig of 3273 bp in length  
41266: 41365: gap of 100 bp  
41366: 52400: contig of 11035 bp in length

FEATURES  
SOURCE

52401 52500: gap of 100 bp  
52501 63506: contig of 11006 bp in length  
63507 63606: gap of 100 bp  
63607 79362: contig of 15756 bp in length  
79363 79462: gap of 100 bp  
79463 97023: contig of 17361 bp in length  
97024 97123: gap of 100 bp  
97124 124426: contig of 27303 bp in length  
124427 124526: gap of 100 bp  
124527 149337: contig of 24811 bp in length  
149338 149437: gap of 100 bp  
149438 181009: contig of 31572 bp in length.

Location/Qualifiers

1..181009  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="RP11-11616"  
/clone\_lib="RP11-11616 Human Male BAC"  
1..1504  
/note="assembly-fragment"  
1605..2791  
/note="assembly-fragment"  
2892..3909  
/note="assembly-fragment"  
4010..5556  
/note="assembly-fragment"  
5557..7482  
/note="assembly-fragment"  
7483..8926  
/note="assembly-fragment"  
8927..10470  
/note="assembly-fragment"  
10571..12199  
/note="assembly-fragment"  
12200..14087  
/note="assembly-fragment"  
14088..15707  
/note="assembly-fragment"  
15708..17945  
/note="assembly-fragment"  
18046..20017  
/note="assembly-fragment"  
20118..22115  
/note="assembly-fragment"  
22216..24556  
/note="assembly-fragment"  
24557..27568  
/note="assembly-fragment"  
27569..32527  
/note="assembly-fragment"  
32528..37892  
/note="assembly-fragment"  
37893..41265  
/note="assembly-fragment"  
41266..52400  
/note="assembly-fragment"  
52401..63506  
/note="assembly-fragment"  
63507..79362  
/note="assembly-fragment"  
79363..97023  
/note="assembly-fragment"  
97024..97123  
/note="assembly-fragment"  
97124..124426  
/note="assembly-fragment"  
124427..149337  
/note="assembly-fragment"  
149338..181009  
/note="assembly-fragment"

BASE COUNT 57610 a 36532 c 33173 g 51292 t 2402 others





```

repeat_region /note="match: GSS: Em:AQ094704"
19777..19993
/note="MER58A repeat: matches 2..224 of consensus"
20817..21173
/note="match: GSS: Em:AQ376698"
complement(20830..20939)
/note="match: GSS: Em:AQ229933"
complement(20830..21133)
/note="match: GSS: Em:AQ235822"
complement(20830..21137)
/note="match: GSS: Em:B65379"
21100..21154
/note="11 copies 5 mer aaac 78% conserved"
21183..21410
/note="MER30 repeat: matches 1..230 of consensus"
21647..21773
/note="MIR repeat: matches 119..246 of consensus"
complement(23139..23500)
/note="match: GSS: Em:AQ210496"
23363..23601
/note="MIR repeat: matches 2..256 of consensus"
complement(24372..24810)
/note="match: GSS: Em:AQ178605"
26871..27202
/note="MSTD repeat: matches 1..341 of consensus"
27203..27695
/note="LIPAI0 repeat: matches 5642..6165 of consensus"
27696..27750
/note="MSTD repeat: matches 341..394 of consensus"
27867..27971
/note="MER5A repeat: matches 84..189 of consensus"
27986..28150
/note="MER5A repeat: matches 3..183 of consensus"
28765..29329
/note="match: GSS: Em:B93338"
28777..29151
/note="match: GSS: Em:AQ125799"
29425..29482
/note="23 copies 2 mer ac 87% conserved"
29740..29805
/note="33 copies 2 mer aa 71% conserved"
29763..29807
/note="9 copies 5 mer aaac 82% conserved"
30205..30412
/note="MIR repeat: matches 2..230 of consensus"
30582..30732
/note="MER5A repeat: matches 8..187 of consensus"
30911..31378
/note="match: GSS: Em:AQ546895"
32009..32080
/note="MIR repeat: matches 132..206 of consensus"
33125..32784
/gene="ba31k16.1"
32125..32784
/gene="ba31k16.1"
/note="ba31k16.1 (snORNA binding domain pseudogene)
match: proteins: Sw:000567 Tr:09VM69 Tr:09ZRM0 Tr:09VDD1
Tr:094514 Sw:021276 Tr:09SG37 Tr:029158 Sw:012460
Tr:080401 Tr:065334 Tr:065335 Tr:09USW4"
/codon_start=1
/pseudo
evidence=not_experimental
34666..34977
/note="AluSg repeat: matches 11..297 of consensus"
36947..36978
/note="16 copies 2 mer aa 87% conserved"
38608..39160
/note="match: GSS: Em:AQ353745"
complement(40339)
/gene="ELAVL2"
complement(40339..43119,43683..43721,51615..51840,
55153..55306,81257..81360,112241..112481,171869..171992))
/gene="ELAVL2"

```

Query Match	Best Local Similarity	Matches	Score	DB %	Length
133: Conservative	4.0%; 64.6%;	0: Mismatches	89.2; 8.9e-14;	73;	183483;
				Indels	0; Gaps
					0
891	ctggagcccaagcttgaggcaactcttcctgtaagaagccagcaaggaaccattttagcttc	950			
10360	CTAGGCACAGAGGTGACGAACTTTCTCTGTTAAAGCCCGAGTGATTAATTTTAACTTT	10419			
951	tgtgtgccatacgtctcagtcacaactactatctctcgtcctctgtagagcaagaagcaat	1010			
10420	TGTGGGCCATATATGCTCTCTGTACACAACATCTCCACTCTCTCTTTCTTAGGGCAAAAGCAAT	10479			
1011	tagcaacaatalgltcaacaacacatalgltgaccoccatgaaacatttattatlatatgatcat	1070			
10480	TATGAGAAATACATTAATGCAATGAGAGATGGCTGTGCCATTAATAAGATTATTATGACAC	10539			
1071	ggaaacctggaataataatgtcttctt	1096			
10540	TGAAATTTGAAATTTGATCATGTAAATTTT	10565			

Search completed: August 16, 2002, 06:44:24  
Job time: 21728 sec



1

2

3

4

5



168th Street, BB 1505, New York, NY 10032, USA  
FEATURES  
Location/Qualifiers  
source 1..2256  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/issue\_type="Placenta"  
promoter 1..2208  
/gene="ST16"  
1..>2256  
/gene="ST16"  
/note="mda-7"  
2209..>2256  
/gene="ST16"  
mRNA  
/product="suppression of tumorigenicity 16 protein"  
BASE COUNT 539 a 584 c 503 g 630 t  
ORIGIN

Query Match 98.6%; Score 2255; DB 9; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 tcaactcttgaaccaggctcgcgcctcccaagcttgtaacttaactagatctcaa 91  
DB 1 tcaacttttgaaccaggctcgcgcctcccaagcttgtaacttaactagatctcaa 60  
QY 92 ctgactgttggccaaaggttccttaagttctctcttgacctctctctgaagtaaatgac 151  
DB 61 ctgactgttggccaaaggttccttaagttctctcttgacctctctctgaagtaaatgac 120  
QY 152 tatgataagctatcagagagcttggagcccaagacatggttggcctgaagatcctca 211  
DB 121 tatgataagctatcagagagcttggagcccaagacatggttggcctgaagatcctca 180  
QY 212 tatgataagctatcagagagcttggagcccaagacatggttggcctgaagatcctca 271  
DB 181 tatgataagctatcagagagcttggagcccaagacatggttggcctgaagatcctca 240  
QY 272 ccaagtgatcaatgaactgaactgaactgttccagcctgccttcaagatcctca 331  
DB 241 ccaagtgatcaatgaactgaactgaactgttccagcctgccttcaagatcctca 300  
QY 332 cctgcctgcctcccgccgacatactgacagagtagtctgaagattcagcctgcctggagg 391  
DB 301 cctgcctgcctcccgccgacatactgacagagtagtctgaagattcagcctgcctggagg 360  
QY 392 ctcaatttcctctcaagctctcgtccttaagctgtctcctgcctcccaactcaattact 451  
DB 361 ctcaatttcctctcaagctctcgtccttaagctgtctcctgcctcccaactcaattact 420  
QY 452 ccagcactcgaactggttct 511  
DB 421 ccagcactcgaactggttct 480  
QY 512 aatagtagttagttaggttct 571  
DB 481 aatagtagttagttaggttct 540  
QY 572 tgggggaagagagctgagcgttggggcgtgagcgttgaagagagagagagagagagag 631  
DB 541 tgggggaagagagctgagcgttggggcgtgagcgttgaagagagagagagagagagag 600  
QY 632 ccaggaacaaagagatcagagctggtgacatataagccaaagcagcgtcgtgccaagagat 691  
DB 601 ccaggaacaaagagatcagagctggtgacatataagccaaagcagcgtcgtgccaagagat 660  
QY 692 tgcagaagagatatttcttct 751  
DB 661 tgcagaagagatatttcttct 720  
QY 752 ggggggttctggtgccaagagagatgggaagagctgccaagagagagagagagagagag 811  
DB 721 ggggggttctggtgccaagagagatgggaagagagagagagagagagagagagagagag 780

QY 812 cctcttct 871  
DB 781 cctcttct 840  
QY 872 ctactccggttcccatatgctcgtgacccaggttgggcaaacctctcctgttaagaagacaga 931  
DB 841 ctactccggttcccatatgctcgtgacccaggttgggcaaacctctcctgttaagaagacaga 900  
QY 932 caggaactattttaggctctgtgtgacatagctcagtaacaaactactatctctctcgc 991  
DB 901 caggaactattttaggctctgtgtgacatagctcagtaacaaactactatctctctcgc 960  
QY 992 tctgtagaagaaagcaatgagcaaatatgttaaaaaataatgtgaccatcagtaaa 1051  
DB 961 tctgtagaagaaagcaatgagcaaatatgttaaaaaataatgtgaccatcagtaaa 1020  
QY 1052 cttattatattatgatacagaaacccgaaataatgtctctctctctctctctctctccca 1111  
DB 1021 cttattatattatgatacagaaacccgaaataatgtctctctctctctctctctctccca 1080  
QY 1112 atcatttaaaaaacgttaaaactactcttaggtgcaaggttgaagcaatctcagctbagc 1171  
DB 1081 atcatttaaaaaacgttaaaactactcttaggtgcaaggttgaagcaatctcagctbagc 1140  
QY 1172 agtggcagagctggaattcgtgcttgagcctacagcttgagcaatccctgattcccaaatgt 1231  
DB 1141 agtggcagagctggaattcgtgcttgagcctacagcttgagcaatccctgattcccaaatgt 1200  
QY 1232 attcctcaagagatgtgggcaaatcttaaggagatgctgatttaaacagagatgaagag 1291  
DB 1201 attcctcaagagatgtgggcaaatcttaaggagatgctgatttaaacagagatgaagag 1260  
QY 1292 catcagaacttccagagagaggttgaagacatgcaagagcctcttaactgaaccatctga 1351  
DB 1261 catcagaacttccagagagaggttgaagacatgcaagagcctcttaactgaaccatctga 1320  
QY 1352 tcaactgttctcaatgagagatctctgcaagaacaaagaaatctcctcaaacctagaagctgagga 1411  
DB 1321 tcaactgttctcaatgagagatctctgcaagaacaaagaaatctcctcaaacctagaagctgagga 1380  
QY 1412 ctgtgcttggggaacacagctctctgcttgatgagcctcaactggtgacatggttagaactcag 1471  
DB 1381 ctgtgcttggggaacacagctctctgcttgatgagcctcaactggtgacatggttagaactcag 1440  
QY 1472 agctgagctgctgcaagctgagatagagctcagagcctctgagcctctgagagctgag 1531  
DB 1441 agctgagctgctgcaagctgagatagagctcagagcctctgagcctctgagagctgag 1500  
QY 1532 aacacatcaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1591  
DB 1501 aacacatcaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1560  
QY 1592 accagagacacagagcctccagctctggtgctgctgctgctgctgctgctgctgctgctg 1651  
DB 1561 accagagacacagagcctccagctctggtgctgctgctgctgctgctgctgctgctgctg 1620  
QY 1652 cctggttccagaaaccttaagaagagacatttggcttgaagatataaacctctggagac 1711  
DB 1621 cctggttccagaaaccttaagaagagacatttggcttgaagatataaacctctggagac 1680  
QY 1712 catggttagctgacatgctgaagaaacatccctcaactctctgcccctgaltgaagaa 1771  
DB 1681 catggttagctgacatgctgaagaaacatccctcaactctctgcccctgaltgaagaa 1740  
QY 1772 gaggatcttaagagttatttctctgctgtgagctactgtcccaagcatgaatgagacatca 1831  
DB 1741 gaggatcttaagagttatttctctgctgtgagctactgtcccaagcatgaatgagacatca 1800  
QY 1832 ttaattgaattatttgaagaaagagagatggtgcatgctgacagtaataatgattttaa 1891  
DB 1801 ttaattgaattatttgaagaaagagagatggtgcatgctgacagtaataatgattttaa 1860

```

-----
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgc@gsu.washington.edu
Drafting Center: SC
-----
Project Information
Center project name: chr-1
Center clone name: RP11-564A8 (SC0742)
-----
Summary Statistics
Sequencing vector: plasmid; 3% of reads
Sequencing vector: plasmid; L08732; 61% of reads
Chemistry: Dye-terminator ET; 8% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 192768 bases at least Q40
Consensus quality: 193102 bases at least Q30
Consensus quality: 193186 bases at least Q20
Insert size: 193217; sum-of-contigs
Quality coverage: 10.0x in Q20 bases; sum-of-contigs
-----

```

	* consists of 2 contigs. The true order of the pieces
	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	1 56155: contig of 56155 bp in length
	* 56156 56255: gap of unknown length
	* 56256 193317: contig of 137062 bp in length.
FEATURES	
source	1..193317
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/clone="RP11-564A8"
	/clone_lib="RPCI human BAC library 11"
misc_feature	1..56155
	/note="assembly_name:Contlg81"
misc_feature	56256..193317
	/note="assembly_name:Contlg82"
BASE COUNT	53599 a 42586 c 43895 g 53137 t 100 others
ORIGIN	
Query Match	96.6%; Score 2208.6; DB 2; Length 193317;
Best Local Similarity	98.8%; Pred. No. 0;
Matches 2253; Conservative	0; Mismatches 9; Indels 19; Gaps 2;
Oy 25	gaccgcgtacccctttgaaccagtgctgcgcgtccccaagaagcttgaactcataactaa 84
Db 103494	GCGCTGATACCCTTTTGAAACCAAGGTCTCCCTGCCTCCAAAGCTGTGAATCATTAAGA 103553
Oy 85	tctcacaactgatgtggcccaagtcttcctaaggltctctctctgaacctctctgaagia 144
Db 103554	TTCACAACGTATGTTGGGCCAAGGTTCCTAGTTCCTCCTTGACACTTCCTTGAAAGTA 103613
Oy 145	ataatgcatagaaaactcatcgaggagctgaagcccagaacacatttgtgccttaactatc 204
Db 103614	ATAATGCTATGATGAAGCTCATTCGGAAGCTGAAGCCCCAAGCACATGTTGGCTGAATTAC 103673
Oy 205	cattgtatatgatctctctctcagaagaagtagactatactcaagatcccaagtgtaacctg 264
Db 103674	CATGTTATATGATATCTCTCTCCCTCGAACAGATGAGCTACTCAAGATCCCAGGTACCTCG 103733
Oy 265	aggccaccaagaagtatcacatbaactatagctcttgftlccaagctgcgccctttaaacggt 324
Db 103734	AGGCCACACCAAGGTATCATATGACCTCAATGCTCTGTTCACAGCTTCGCCCTTTAAACGCT 103793
Oy 325	catccacacgcgcgtgcctcccccgcctacatcgcagaacagtagctatgaatttaagctgcc 384
Db 103794	CATCCACACGCGCTGCCTCCCTCCGCCCTATCTCAACAGTAGTGTAGGATTTTCAAGCTGCC 103853
Oy 385	ctggaggcctaatttcocctcctcagaagctccctgtcttaagctatgcctctcctccccaactacc 444
Db 103854	CTGGGGGCTCATTTTCCCTCTCAAGCTTCCCTGCTTTAAGCTGTCTCTGCTCCACATCCAC 103913
Oy 445	tattactccagcaactcctaacctggtctctcttllctgtctcaalcacgctcttgacaactc 504
Db 103914	TATTACTCCAGCACTCTCACTCTGCTGTTCTTTCTCTCAATCACTGCTCTTGCATCACTTC 103973
Oy 505	ttatctcatagtagttagtaaggaggtctctgtgtaatgccccaaatccacatggtgtggaag 564
Db 103974	TTATCTCATAGTACTGTTAGTTAAGGGGTTCTTGATATGCCCTTAATCCACATGATGATGAG 104033
Oy 565	ggggagatbgggagaaagagtagtcgctgtvgggcgtgtgcctactctcggagagtgtaagactc 624
Db 104034	GGGGAGATBGGGAGAAAGAGTAGTCTGCTGTGGGGCTGTGCCCTTCTTGAGAGGTAAAGCTATC 104093
Oy 625	ggggccctccaggaacaagaattccagctcgtgtgggaagctatagccaagaagactcgtcggc 684
Db 104094	GGGGCCCTCCAGGAACAAGATTTCAAGCTGTGGGCGACGCTATATAGCCMAAGCACTGCTGGCC 104153

OY	685	caggagatgtaaaaggagatcttctgttgccttaaaataaaacaacagatagat	744
Db	104154	CAGGAGATTGCAAAAGAGATTTTGTGTTCTTAAACAAATTAACACACTGATATGAKAT	104213
OY	745	ggagaggagggtgtgtgtgcagagagattggaagagctgtccaaagtgtgtctac	804
Db	104214	GGAGGGAGGGGGCTGTGGTCCAGAGATTTGGGAAGCTCTCCAAAGGTTGTTCAC	104273
OY	805	tcactccctcttcttctcaatctcaatgtagctggaggagtataccgtccccaagt	864
Db	104274	TCACTCCCTCTTCTTCTTCACTCCACTGAGCTGGAGGAGATTACCTGTCTCCACGT	104333
OY	865	cacatctcctatcccglttcccaatgtccatgagccaggttggccaactcttcgttaag	924
Db	104334	CACATTCCTAATCTCCGTTTCCATGCCCTGGAGCCAGATTGGGAACCTTCTGTATAAG	104393
OY	925	aaccagagagaaacttcttaagctctgtgtccatatgttccagtcacaatactcat	984
Db	104394	AACCAGAGAGAACTATTATTAGCTCTGTGTGCATATGTGTCACTCAACTACTCAT	104453
OY	985	ctctgcctctgttaacacgaagcaattagcaacaatatttcaacaacatagtaccac	1044
Db	104454	CTCTGCCTCTGTATGACAGAAACCAATTATGCAACATATTGTCAACCAACTATATGACCCC	104513
OY	1045	atgaaaaacttacttacttgatacgaaacccgaaataatgtcttcttcttattt	1104
Db	104514	ATGAAAACTTATTTATTATGTGATACGAAACCTGAAATAATATGTCTTCTTATTTT	104573
OY	1105	ttccccaatctatataaaaaagaaaacctctctgtgtcgcagaagttaagcatctca	1164
Db	104574	TTCCCCAATCTATTAAAAACGTAATAACTCTTAAAGTCGCAAGGTTAAGCAATTCTCA	104633
OY	1165	gcttagcagtgagagcttgatcttggtctgtgatacaglttgccaatccctgattccc	1224
Db	104634	GCTTAGCAGTGGAGGCTGATTTGGCTTGACACTACATTGGCCAACTCCGTATTTCC	104693
OY	1225	aaaatctatccctcagaggaatggggcaaatatctatgtggaaagtgcctgataaacaagt	1284
Db	104694	AAAATCTATTCCTCAAGGATGTGGCAAAATCTTATGGAATAAGCTGGATTAAACAGAT	104753
OY	1285	taagaagcatcagaacatctccaagacggtctagacatgacagagctctctaactgact	1344
Db	104754	TAAGAAGCATTCAGACATTTCCAGAGCGGCTAGCACATCCAGAGGCTCTTAACGACCT	104813
OY	1345	caatcggatctatcgttctcatgtgaagatcttcagaaacaataatctccaactagag	1403
Db	104814	CATTAGCATTCATCTGTTTCATATGAGATCTTGCACAGACAGAAATTCCTCAACCTTAAG	104873
OY	1404	tcctgagacgtgtgcttcggaaacaatgtctgtctgaatgcccctcaactgtggacaatgta	1463
Db	104874	TCCTGAGACTGTGCTTTGGGAACACTGCTCTGTGATGCCCTCACTGGGACATGGTTA	104933
OY	1464	gaatctagagcttgagtgctcttctcagctcgagaaataagagtcdaagctcttgaactccctgg	1523
Db	104934	GAACTTAAGACTGAGTGCCTTCCTGATAGCTGGAAGATAGGCTTCAGAGCTTTGACTGCCCTGG	104993
OY	1524	cagctctgaacaacaacgcgtgtctgtgtccctcgaatgtgttcaagacacaagaccag	1583
Db	104994	CAGCTTTGACACATCAACCTGCTGTGTGCTCCCTAGTGGTTCAAGACCAACAGGCCAAG	105053
OY	1584	actagccacaacagagacacagagcctcccaactcttcgggctgtgtccatgtgtatacttc	1643
Db	105054	ACTAGCCACACAGACACACAGGCTCCCACTTTCTGGGCTGTGCAATGCTACTATTTCC	105113
OY	1644	tcatctctccctgtgttccagaacctaaaggagagcaatctgtgttagtgatataaac	1703
Db	105114	TTCATCTCTCTCGTTTCCAGAACTTAAGAGAGGACACTTTGGTTGATATTAATTAAC	105173
OY	1704	ctagggacaatgggttagctgtcaatcttaaggaaacaaccccaactctccgtggccctgattgga	1763
Db	105174	CTAGGACCAATGGGTAGCTGCATGTCAAGAAACACTCTCAACTTCTGTGCGCCCTGATATGA	105233
OY	1764	ttaaaaggagaggtactcaaggttaattctctgcgtgtgatactatgttcccaatgata	1823

Db	105234	TTAAAGAGAGGCTACTACAGTTATTTCTTCGCTGTGACACTACTGTCCAGCATGATA	105293
QY	1824	gggcacatcatttgaattatttgcaggaagagacatggtatgctgcacgtaaa	1883
Db	105234	GGGCGATATATATGAAATATTTTGTGACAGAGAGAGACTGGTGTATGCGACACTAATAA	105353
QY	1884	tgtattacaigtgtacagagtttaaccaagacactctgtctgttcttgccttgttat	1943
Db	105344	TGTTTACATGTGTGACAGATTACCAAGACACTCTGTGTGTTTGGCCATGTTTAT	105413
QY	1944	tacacct-----gggacaaattttaaattatcatagagaaact	1985
Db	105414	TACACTTGCCCTTGGGGGAGGTGGACAAATTTTAAAAATTTTATCATGCGAGAACT	105473
QY	1986	gcagcgagagaagactlaagagactgtccctctgccacaacagcagtgtagacctgaac	2045
Db	105474	GCACCGCAGAGAAAGCTAAAGACTTGGCCCCCTGCCACACAGCACTGGTATGAGCCTGAC	105533
QY	2046	tcaaaccaagcttcaatccacctaaagggtgtcttccocactgcgtatctgacctaa	2105
Db	105534	TCAAACCCAGCTCTACTCTCACTCTAGGGGTGGTTTCCGCAATCGCTGTATTTGTCCTAA	105593
QY	2106	agtgtatgtgtgacttagcaatgaagtaattctctctaggaaagcagaaccaatttcccttc	2165
Db	105594	AGTATGTGGTACTAGGCAATTAAGTAATTCCTACGAAAGCATCAACAAATTTCCCTTTC	105653
QY	2166	tccacctccctcttcttccccaacccctccccaatcagcccccaatalatgaccaaatc	2225
Db	105654	TCCACCTCCCTCTTTTCTCTCCACCCCTCCCATCAAGCCCCCAATATATATGSCCAATC	105713
QY	2226	tcccaagaagctgtctgtgcgcaaacctttaacttcttgaaatgaacttccaaggtctgac	2285
Db	105714	TCACCAAGCCTTGCTGTGCTCTCCAAACCTTTACTCTTGAAATGACTTCCACAGGCTGGAC	105773
QY	2286	g 2286	
Db	105774	G 105774	

AC023534/C LOCUS	RESULT 3			
DEFINITION	AC023534	229302 bp	DNA	linear
ACCESSION	Homo sapiens clone RP11-462N18	WORKING DRAFT	SEQUENCE	41
VERSION	AC023534			
KEYWORDS	uncloned pieces.			
SOURCE	AC023534.3	GI:7705140		
ORGANISM	HGTG; HGTG_PHASE1; HGTG_DRAFT.			
REFERENCE	human.			
AUTHORS	Homo sapiens			
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 229302)			
REFERENCE	Birren, B., Linton, L., Nusbaum, C. and Lander, E.			
AUTHORS	Homo sapiens chromosome, clone RP11-462N18			
	Unpublished			
	2 (bases 1 to 229302)			
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,			
	Anderson, S., Balgwin, J., Barra, N., Beld, F., Boguslavsky, L.,			
	Boukhalgal, B., Brown, A., Burkett, G., Campolino, A., Castle, A.,			
	Choegel, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, P.,			
	Dearlano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,			
	Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,			
	Galaan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,			
	Grand, P., Grant, G., Hagos, B., Heatford, A., Horton, L.,			
	Howard, J., Iliev, I., Johnson, R., Jones, C., Kani, L., Karatas, A.,			
	Klein, J., Landers, T., Larocque, K., Lebowitz, J., Levine, R.,			
	Lieu, C., Liu, C., Locke, K., MacDonald, P., Marquis, N., McCarthy, M.,			
	McBain, P., McQuar, A., McKernan, K., McMichael, R., Meldrum, J.,			
	Menais, L., Mitova, T., Miranda, C., Mieng, V., Morrow, J., Naylor, J.,			
	Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M.,			
	Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C.,			
	Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,			



TITLE  
JOURNAL

## COMMENT

Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Subramanian, A., Talamas, J., Testaye, S., Theodode, J., Tirrell, A.,  
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J., Zimmer, A. and  
Zody, M.

## Direct Submission

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 330 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2000 this sequence version replaced gi:7143452.  
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

## Project Information

Center project name: L5244

Center clone name: 462\_M18

## Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 194690 bases at least Q40

Consensus quality: 207622 bases at least Q30

Consensus quality: 217061 bases at least Q20

Insert size: 200000; agarose-gel

Insert size: 225302; sum-of-coverage

Quality coverage: 4.5 in Q20 bases; agarose-gel

Quality coverage: 4.0 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently  
consists of 41 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1108: contig of 1108 bp in length  
\* 1109 1208: gap of 100 bp  
\* 1209 2504: contig of 1296 bp in length  
\* 2505 2604: gap of 100 bp  
\* 2605 3688: contig of 1084 bp in length  
\* 3689 3788: gap of 100 bp  
\* 3789 4962: contig of 1174 bp in length  
\* 4963 5062: gap of 100 bp  
\* 5063 6475: contig of 1313 bp in length  
\* 6476 7673: contig of 1198 bp in length  
\* 7674 7773: gap of 100 bp  
\* 7774 9425: contig of 1651 bp in length  
\* 9426 9525: gap of 100 bp  
\* 9526 10997: contig of 1472 bp in length  
\* 10998 11097: gap of 100 bp  
\* 11098 12740: contig of 1643 bp in length  
\* 12741 12840: gap of 100 bp  
\* 12841 14467: contig of 1627 bp in length  
\* 14468 14567: gap of 100 bp  
\* 14568 16980: contig of 2413 bp in length  
\* 16981 17080: gap of 100 bp  
\* 17081 19176: contig of 2096 bp in length  
\* 19177 19276: gap of 100 bp  
\* 19277 21882: contig of 2606 bp in length  
\* 21883 21982: gap of 100 bp  
\* 21983 23826: contig of 1844 bp in length  
\* 23827 23926: gap of 100 bp  
\* 23927 26561: contig of 2635 bp in length  
\* 26562 26661: gap of 100 bp  
\* 26662 28846: contig of 2185 bp in length  
\* 28847 28946: gap of 100 bp  
\* 28947 32495: contig of 3549 bp in length

## FEATURES

## source

32496 32595: gap of 100 bp  
\* 32596 34950: contig of 2355 bp in length  
\* 34951 35050: gap of 100 bp  
\* 35051 38193: contig of 3143 bp in length  
\* 38194 38293: gap of 100 bp  
\* 38294 41862: contig of 3569 bp in length  
\* 41863 41962: gap of 100 bp  
\* 41963 45898: contig of 3936 bp in length  
\* 45899 45998: gap of 100 bp  
\* 45999 50904: contig of 4906 bp in length  
\* 50905 51004: gap of 100 bp  
\* 51005 55346: contig of 4342 bp in length  
\* 55347 55446: gap of 100 bp  
\* 55447 61337: contig of 5890 bp in length  
\* 61337 61436: gap of 100 bp  
\* 61437 65806: contig of 4370 bp in length  
\* 65807 65906: gap of 100 bp  
\* 65907 71648: contig of 5742 bp in length  
\* 71649 71748: gap of 100 bp  
\* 71749 74886: contig of 3138 bp in length  
\* 74887 74986: gap of 100 bp  
\* 74987 79878: contig of 4892 bp in length  
\* 79879 79978: gap of 100 bp  
\* 79979 84280: contig of 4302 bp in length  
\* 84281 84380: gap of 100 bp  
\* 84381 89873: contig of 5493 bp in length  
\* 89874 89973: gap of 100 bp  
\* 89974 97168: contig of 7195 bp in length  
\* 97169 97268: gap of 100 bp  
\* 97269 101906: contig of 4638 bp in length  
\* 101907 102006: gap of 100 bp  
\* 102007 109439: contig of 7433 bp in length  
\* 109440 109539: gap of 100 bp  
\* 109540 117592: contig of 8053 bp in length  
\* 117593 117692: gap of 100 bp  
\* 117693 126037: contig of 8345 bp in length  
\* 126038 126137: gap of 100 bp  
\* 126138 135140: contig of 9003 bp in length  
\* 135141 135240: gap of 100 bp  
\* 135241 146384: contig of 11144 bp in length  
\* 146385 146484: gap of 100 bp  
\* 146485 160352: contig of 13868 bp in length  
\* 160353 160452: gap of 100 bp  
\* 160453 173458: contig of 13006 bp in length  
\* 173459 173558: gap of 100 bp  
\* 173559 191705: contig of 18047 bp in length  
\* 191706 229302: contig of 37597 bp in length.  
Location/Qualifiers  
1..229302  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="RP11-462N18"  
/clone\_id="RPCT-11 Human Male BAC"  
1..1108  
/note="assembly\_fragment"  
1209..2504  
/note="assembly\_fragment"  
2605..3688  
/note="assembly\_fragment"  
3789..4962  
/note="assembly\_fragment"  
5063..6475  
/note="assembly\_fragment"  
6476..7673  
/note="assembly\_fragment"  
7674..9425  
/note="assembly\_fragment"  
9426..10997  
/note="assembly\_fragment"  
11098..12740  
/note="assembly\_fragment"  
12841..14467

```
/note="assembly-fragment"
misc_feature 14568..16980
/note="assembly-fragment"
misc_feature 17081..19176
/note="assembly-fragment"
misc_feature 19277..21882
/note="assembly-fragment"
misc_feature 21883..23826
/note="assembly-fragment"
misc_feature 23927..26361
/note="assembly-fragment"
misc_feature 26662..28846
/note="assembly-fragment"
misc_feature 28947..32495
/note="assembly-fragment"
misc_feature 32596..34950
/note="assembly-fragment"
misc_feature 35051..38193
/note="assembly-fragment"
misc_feature 38294..41862
/note="assembly-fragment"
misc_feature 41963..45898
/note="assembly-fragment"

Query Match 96.6%; Score 2208.6; DB 2: Length 229302;
Best local similarity 98.8%; Pred No. 0;
Matches 2253; Conservative 0; Mismatches 9; Indels 19; Gaps 2;

Oy 25 gactcgtacaccccttgaacccagctcgtcgtcccaaaagctgtacataactaga 84
DB 158375 GCGTCGATGACACCTTTTAAACCCAGGCTGCTGCTCCCAAACTTGATATACAGA 158316
Oy 85 ttcccaactgaagtgctgggccaaggctcccaaggtctctctctgaacccctcctgaagta 144
DB 158315 TTCTCAACTGATGTTGGGGCCAAAGTCTAGGTTCTCTCTGACCTTCTTGAGTA 158256
Oy 145 ataattgataataagctacagagctgagagcccaagccacatgttctgctgaacatc 204
DB 158325 ATATGCTATGATAGTCACTGATCGAGGCTGAGCCAGGACATTTTGGCTGAACTATC 158196
Oy 205 cagtgtatatagaatcctctctcctcagacagagtgagctacatcagatcccaagtgtaacctg 264
DB 158195 CATGTTATGATGATCTCTCTCAAGAGAGTACGATCAACAGATCCAGGCTGTAACCTG 158136
Oy 265 aggcacagccaagtgatcatcactcaagctcgtctcgtctcagcctgacctttaaagct 324
DB 158135 AGGCACAGCAAGGTGATCATGACTCATGCTCTGTTCCAGCCTTGCCCTTAAAGCT 158076
Oy 325 catcccaacctgctcgtcctcccgctacatcgcagacagtagtctaggaattcagctgccc 384
DB 158075 CATCCACACTGCTGCTGCTCCCGCTATCTCAAGACATGATGATTAAGATTTCAGCTGCC 158016
Oy 385 ctggggagctcaattctcctctcagctcctcgtctctctgaagctcctcgtcctcctcaac 444
DB 158015 CTGGGGGCTCATTTTCCCTCTCAGCTTCCGTTAGCTGCTCCTGCTGCCACTCACTACC 157956
Oy 445 tattaacctcggacctcctcaacctgtcttcttctgtctctatactgaactgtccttcaactc 504
DB 157955 TATTACTCAGCACTCTCAGCTGCTGCTCTCTTCTCTCTCACTGCTGCTTAAACATCT 157896
Oy 505 ttatctcatagtagttagttaggggtctctgttaaagccttaaatccatctggtgggaag 564
DB 157895 TTATCTCATGATGATGATGATGAGGGGTTCTTGTAATGCCCTTAATCCACATGGGGGAG 157836
Oy 565 gggggagtgagggaaggaagtgctgtggggctgtgacctactcttgaaggttaagctc 624
DB 157835 GGGGGAATGGGGGAAGAGAGTGGCTGTGGGGCTGTGGCTACTCTGTGAAGGGTAAGACTC 157776
Oy 625 gggccctccaggaacaaagatcaggctgtgtgagcgtatagccaagcagaactgctgyc 684
DB 157775 GGGCCCTCCAGGAACAAAGATTCAAGGCTGTGGCAGCATATAGCAAGCAACAGCTGTGCGC 157716
Oy 685 cagggtctgcaagagtagtattgtctgtcttaagaaataaacaacctgtagcttagat 744

|||||
DB 157715 CAGGAGATTGCNAAGGAGATATTGTTGCTTAAGAAATAAACAACCTGATGATGAT 157656
Oy 745 ggaaggagggggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 804
DB 157655 GGAGGAGAGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157596
Oy 805 tcaactcctcctctctctctctctctctctctctctctctctctctctctctctctctct 864
DB 157595 TCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 157536
Oy 865 cacattcctaccccgcttcccaagcctgtgagccaggtgtgggcaaacctcctcctgaag 924
DB 157535 CACATTCCTCACTCCGCTTCCGCTGCTGAGCCAGCTGAGCCAGCTGAGCCAGCTGAGCCAG 157476
Oy 925 aaccagagcaggaacttttaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 984
DB 157475 AACCAGAGGAGAACTATTATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157416
Oy 985 ctctgctctctgttagcaagaaagcaatcagcaacaaatagcaacaaatagcaacaaatagca 1044
DB 157415 CTCTGCTCTCTGTAGCAAGAAAGCAATTAACCAATATGTCACAAACATATGTGACCC 157356
Oy 1045 atgaaacttatatttatatgtatagacgaaacctgaataatgtctctctctctctctct 1104
DB 157355 ATGAAACCTTATTATTATGATGATGATGATGATGATGATGATGATGATGATGAT 157296
Oy 1105 ttccccaatcaataaataaataaataaataaataaataaataaataaataaataaataaataa 1164
DB 157295 TTCCCAATCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 157236
Oy 1165 gctcagagctgagagcctgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1224
DB 157235 GCTTACGAGTGTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157176
Oy 1225 aaatgfatctcccaagagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1284
DB 157175 AAATGATATCTCTCAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157116
Oy 1285 taagaagcatcaagacattccagagcgtgtagcaatcagcagcgtctcctcaactgaact 1344
DB 157115 TAAAGACATCAAGACATTTCACAGAGAGGCTGTGACATGACCTTCACTGACCT 157056
Oy 1345 caat-aggatcatcctgtctcctcagagagctgtgagagcaagatcctcaaacctagag 1403
DB 157055 CATTAAGATTCACTGCTGTTTCATGAGAGATCTGCAAGCAAGATTCCTCAAACTAGAG 156996
Oy 1404 tctgagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1463
DB 156995 TCTGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156936
Oy 1464 gaactcagagctgagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1523
DB 156935 GAATCTAGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156876
Oy 1524 cagctctgagacatcagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1583
DB 156875 CAGCTTGAGACATCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156816
Oy 1584 actagccacacagagacagagcctcctcagccttctgtgtgtgtgtgtgtgtgtgtgtgtgt 1643
DB 156815 ACTAGCCACACAGAGACAGAGCCTTCCAGCTTCTGAGCTTGTGCTGCTGCTGCTGCTG 156756
Oy 1644 ttattctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1703
DB 156755 TTATTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156696
Oy 1704 ctgaagacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1763
DB 156695 CTAGGACCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156636
Oy 1764 ttaagagagagtagttaggttattctctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1823
```

	Db	156335	TTAAGAGAGGACTTACAGATTATTTCTCGCTGGAGCTACGTCCACAGATGAATA	156576
Oy	1824		ggagcaccattatggaatttatlttgcacagaaggagactgtgatctcgaacgaataa	1883
Db	156575		GGGCATCATTTATTGAATTTTTTCACAGAGAGAGACTGCTGATGTGCACATATAA	156516
Oy	1884		tgtattacaatgtylacagagttcacaaagcactgtgtgttlltgcccttgtaac	1943
Db	156515		TGTATTTCATCATGTGTACAGAGTTTACCAGAGCCTCTGTGTGTATTTTGCTCTGTTTAF	156456
Oy	1944		tacacct-----ggacaaactttcaaatltaatlacatgcagagact	1985
Db	156455		TACACTTGGCCCTTGGGGGAGGTGGGACAATTTTTTAAAAATTATATCATGCAGAGACT	156396
Oy	1386		gcagcccgagaggaagctaaagaagctgtcccctgtgccacagagcaugyttagagcctgaac	2045
Db	156395		GCGACCGAGAAGAGCTTAAGAGACTGCCCTGCCCAACAGCCAGTGTGTAGACCTGAAC	156336
Oy	2046		tcaaaccaggttcatactcaacctcaaggagctgtcttcccatagtgtatltgtccttaa	2105
Db	156335		TCAAACCCAGAGCTCATCTCACCTCAGGGGCTGCTTCCCACATGCGTATTCCTTAA	156276
Oy	2106		agtgatgagtgactgaagcaatgaattaftctcaggaaaagcatgaccaattcccttc	2165
Db	156275		AGTGATGAGTGAGCATAGCAGCAATGAATATTCTTAGAAGAAAGCATGACCAATTTCCCTTTC	156216
Oy	2166		tcaacctccacttlttctctccaacctcccccaatcagccccatatataatgcccaaac	2225
Db	156215		TCCACCTCCCCCTTTTCTCTCCACCCCTCCCAATCAGCCCCCATATATATVGGCCAAAATC	156156
Oy	2226		tccacaagaagcctgtgctgtcgtcgaaccttactctcgaatgactccagcagctgagac	2285
Db	156155		TCCACAAGAGCTGTGCTTCGCTGCANAACCTTACTTGAAATGACTTCCAGCGCTGGAC	156096
Oy	2286		g 2286	
Db	156095		G 156095	

  

RESULT	4
LOCUS	AC068122 180048 bp DNA linear HTG 07-JUL-2000
DEFINITION	Homo sapiens chromosome 1 clone RP11-237c22, WORKING DRAFT
ACCESSION	AC068122
VERSION	AC068122.3 GI:79233971
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 180048)
JOURNAL	Waterston,R.H.
REFERENCE	The sequence of Homo sapiens clone
AUTHORS	Unpublished
TITLE	2 (bases 1 to 180048)
JOURNAL	Waterston,R.H.
REFERENCE	Direct Submission
TITLE	Submitted (28-APR-2000) Genome Sequencing Center, Washington
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
	MO 63108, USA
COMMENT	On May 18, 2000 this sequence version replaced gi:7771167.

  

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0237C22
----- Summary Statistics -----
Sequencing vector: M13: 100%
Chemistry: Dye-primer ET; 100% of reads

```

Chemistry	Dye-terminator	Big Dye	% of reads
Assembly program:	Phrap	Assembly	0.990319
Consensus quality:	158944 bases at least	Q40	
Consensus quality:	165826 bases at least	Q30	
Consensus quality:	169627 bases at least	Q20	
Insert size:	167000;	agarose-fp	
Quality coverage:	3.90 in Q20 bases;	sum-of-contigs	
-----			
* NOTE: This is a 'working draft' sequence. It currently			
* consists of 28 contigs. The true order of the pieces			
* is not known and their order in this sequence record is			
* arbitrary. Gaps between the contigs are represented as			
* runs of N, but the exact sizes of the gaps are unknown.			
* This record will be updated with the finished sequence			
* as soon as it is available and the accession number will			
* be preserved.			
1	1400:	contig of 1400 bp in length	
1401	1500:	gap of unknown length	
1501	3388:	contig of 1888 bp in length	
3388	3484:	gap of unknown length	
3389	4879:	contig of 1391 bp in length	
4880	4979:	gap of unknown length	
4980	6565:	contig of 1986 bp in length	
6566	7065:	gap of unknown length	
7066	9366:	contig of 2301 bp in length	
9367	9467:	gap of unknown length	
9467	11957:	contig of 2491 bp in length	
11958	12057:	gap of unknown length	
12058	14208:	contig of 2151 bp in length	
14209	14308:	gap of unknown length	
14309	17022:	contig of 2714 bp in length	
17123	17122:	gap of unknown length	
17123	20061:	contig of 2939 bp in length	
20062	20161:	gap of unknown length	
20162	24025:	contig of 3864 bp in length	
24026	24125:	gap of unknown length	
24126	26125:	gap of unknown length	
26126	26818:	gap of unknown length	
26819	28618:	contig of 2593 bp in length	
28619	29036:	gap of unknown length	
29037	32937:	contig of 3755 bp in length	
32937	32891:	gap of unknown length	
32892	38895:	contig of 6004 bp in length	
38895	38895:	gap of unknown length	
38996	44306:	contig of 5311 bp in length	
44307	44306:	gap of unknown length	
44407	51589:	contig of 7183 bp in length	
51590	51589:	gap of unknown length	
51890	58551:	contig of 6502 bp in length	
58592	58691:	gap of unknown length	
58692	63442:	contig of 4751 bp in length	
63443	63542:	gap of unknown length	
63543	70554:	contig of 7012 bp in length	
70555	70654:	gap of unknown length	
70655	77794:	contig of 7140 bp in length	
77795	77894:	gap of unknown length	
77895	85841:	contig of 7947 bp in length	
85842	93227:	gap of unknown length	
93228	93327:	gap of unknown length	
93328	100871:	contig of 7544 bp in length	
100872	100971:	gap of unknown length	
100972	11241:	contig of 11170 bp in length	
11242	11241:	gap of unknown length	
11243	126358:	contig of 14117 bp in length	
126359	126458:	gap of unknown length	
126459	141767:	contig of 15309 bp in length	
141768	141867:	gap of unknown length	
141868	155927:	contig of 14060 bp in length	
155928	156027:	gap of unknown length	
156028	180048:	contig of 24021 bp in length.	



```

QY 1984 ctgcagcgacagagactgacctgcccacacagcagtgtagagcctga 2043
    |||
Db 82683 ctgcagcgacagagactgacctgcccacacagcagtgtagagcctga 82742
    |||
QY 2044 actcaaacacagcgtcatcaccatcagggcgcttccaccatgctatgctcct 2103
    |||
Db 82743 actcaaacacagcgtcatcaccatcagggcgcttccaccatgctatgctcct 82802
    |||
QY 2104 aagatgatgggtgactagcgaatgaattctctcagaagaagcctgaccattcctt 2163
    |||
Db 82803 aagatgatgggtgactagcgaatgaattctctcagaagaagcctgaccattcctt 82862
    |||
QY 2164 tctccacctcctcttcttccaccctcccccacagcccatatatagcacaa 2223
    |||
Db 82863 tctccacctcctcttcttccaccctcccccacagcccatatatagcacaa 82922
    |||
QY 2224 tctccacaagcctgctgctgctcacaaccttactctcgaatgactccagcgctgg 2283
    |||
Db 82923 tctccacaagcctgctgctgctcacaaccttactctcgaatgactccagcgctgg 82982
    |||
QY 2284 acg 2286
    |||
Db 82983 acg 82985
    |||

RESULT 5
LOCUS AY062931 11114 bp DNA linear PRI 26-NOV-2001
DEFINITION Homo sapiens Interleukin 24 (IL24) gene, complete cds.
ACCESSION AY062931
VERSION AY062931.1 GI:17104816
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Rider,M.J., Armel,T.Z., Carrington,D.P., Chung,M.-W., Lee,K.L.,
Poel,C.L., Toth,E.J., Yi,Q. and Nickerson,D.A.
TITLE Direct Submission (13-NOV-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattleSNPs, NHLBI Program for
Genomic Applications, UW-FHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu).
FEATURES
Source 1. 11114
Location/Qualifiers
repeat_region
/db_xref="taxon:9606"
/rpt_family="MER1_type"
/rpt_type="dispersed"
variation
/frequency="0.01"
/replacement="a"
variation
/frequency="0.10"
/replacement="c"
variation
/frequency="0.05"
/replacement="g"
repeat_region
/rpt_family="MER1_type"
/rpt_type="dispersed"
variation
/frequency="0.07"
/replacement="a"
variation
/frequency="0.07"
/replacement="a"
variation
/frequency="0.05"
/replacement="g"
repeat_region
/rpt_family="SINE"
/rpt_type="dispersed"
variation
/frequency="0.02"
/replacement="c"

variation 1889
/frequency="0.06"
/replacement="g"
join(1897..2068,2184..2329,3775..3970,4751..4813,
5950..6108,6454..6528,7432..8320)
/gene="IL24"
/product="Interleukin 24"
1897..8320
/gene="IL24"
2049
/gene="IL24"
/frequency="0.06"
/replacement="t"
join(2286..2329,3775..3970,4751..4813,5950..6108,
6454..6528,7432..7515)
/gene="IL24"
/codon_start=1
/product="Interleukin 24"
/protein_id="AA134146.1"
/db_xref="GI:17104817"
/translation="MNFQRLQSLMTLARPFCPLATFASQMMNVIPCLGFTLLMS
QVSGAQGEFHEGPGVQKGVLPQKLMEAFMAVKQFMOAQDNITSAFLQGEVQNVSD
AESCYLVHTLEFLYLTWFKNHNRTVRLKSFSLANFVLIVSQLOPQSENMF
SIRDSARHREFLFRRAFKQLDVEALTRALGEVDILITMQKFKYL"
2324..2640
/rpt_family="SINE"
/rpt_type="dispersed"
2461
/gene="IL24"
/frequency="0.90"
/replacement="a"
2546
/gene="IL24"
/frequency="0.01"
/replacement="a"
2625..3310
/gene="IL24"
/note="Region not scanned for variation"
3221
/gene="IL24"
/frequency="0.01"
/replacement="c"
3260
/gene="IL24"
/frequency="0.37"
/replacement="t"
3292
/gene="IL24"
/frequency="0.01"
/replacement="a"
3352
/gene="IL24"
/frequency="0.01"
/replacement="t"
3466
/gene="IL24"
/frequency="0.01"
/replacement="a"
3517
/gene="IL24"
/frequency="0.02"
/replacement="a"
3530
/gene="IL24"
/frequency="0.05"
/replacement="a"
3531
/gene="IL24"
/frequency="0.99"
/replacement="a"
3625
/gene="IL24"
/frequency="0.37"

```



```

QY 1263 gaatgtgattaaacagagatgaagacatcacatctccagagcggcagacat 1322
DB 901 GAATGTGATTAAACAGAGATTGAAGACATCAACATCTCCAGAGAGCGGCTAGCAT 960
QY 1323 gcaagagctctcaactgacccatc-vgatcaccctgttcaatggagatcttgcaaga 1381
DB 961 GCCAGGCGCTCTTAACTGACCTCATATTAGATTCATCTTTTCATGAGAGATCTTGCAAGA 1020
QY 1382 caagaattcccaaacctagaagcagagagctgtcttggaagaaacctgtcctgtga 1441
DB 1021 CAAGAATTCTCTCAACCTAGAGCTGAGAGAGCTGCTTTGGGAAACACTGCTGCTTGA 1080
QY 1442 tgcctcaactgagacatgttgatgaatcagagctgagtgctgtcagctggagatggg 1501
DB 1081 TGCCCTCACTGGGACATGCTAGAACTTAGAGCTAGGCTCTCTACCTGGAATGAGG 1140
QY 1502 tcaagactctgactgctgctgagctgttgacacacacagctgtcgtcctcctgagt 1561
DB 1141 TCAGAGCTCTTGACTGCTGCGAGCTTGACATCAACGCTGCTGCTGCTGAGTG 1200
QY 1562 gtccagagcacaagagcacaagatagccacacagagcagcagcctccagcttcttg 1621
DB 1201 GTTCAGAGCCACACAGGCGCAAGCTAGCCACAGAGCAGAGGCTCCAGCTTTCTGG 1260
QY 1622 gctgtcactgagctacatttcttcttctcctgtgttccagaaactaagagagcaca 1681
DB 1261 GCTTGTGCTAGCTGATTCCTTCTCTCTGCTTCCAGAACTTAGAGAGAGCACA 1320
QY 1682 tttgtgtgagtgatataacccctagagacatgtgtcgtcgtatgtcagaaacctcc 1741
DB 1321 TTTGTGAGTGAATTAATTAACCTTAGGAGACATGGGTGAGCTGCAAGTCAAGAACTCC 1380
QY 1742 tcaactccctgagcctgagatgaatgaagagagagcttaccagcttcttctgctgtg 1801
DB 1381 TCAACTCTCTGAGCCCTGATGAGATTAAAGAGAGAGTACTACAGCTTATCTCTGCTGTG 1440
QY 1802 gacacagctccagcagatgagatgagagcactatattgaattttttgagagaaagagact 1861
DB 1441 GACTACTGCTCCAGCATGANTAGGGCATATTATTGAAATTTTTCAGAGAGAGAGACTC 1500
QY 1862 ggtgtatgctgacagatgaatgaatgtatatacgtgtacagaggttaccagacactctg 1921
DB 1501 GGTGTATGCTGACAGATTAATGATTTACATGTGTACAGAGTGTACCAAGCAGCTCTG 1560
QY 1922 tttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1981
DB 1561 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
QY 1981 aaaaattatacatgacagagactgacagcagcagagagacttgacccctgcccaca 2033
DB 1621 AAAATTATATACATGACAGAGACTGACAGGAGAGAGCTAGAGACTTCCCTGCCCCA 1680
QY 2033 cagcagatgtagagagctgaactcaaacagagctgctcactcactcagagagagctgtctc 2083
DB 1681 CAGCAGAGGAGAGAGCTGACATCAACCCAGAGTCTCATCTCAGGAGGAGGCTGCTTC 1740
QY 2084 cccatgctgtatgtctcctaaagtgatggtgactgagcagaaagaaatgaatctcctaga 2143
DB 1741 CCCATGCTGTATGTCTTAAGATGATGAGTACAGCATGAAGTAATTTCTCTAGA 1800
QY 2144 aagatgacacatttcccttctccacctcctcttcttccctccacccctccacatcag 2203
DB 1801 AAGATGACCAATTTCCCTTCTCCACCTCCCTTCTTCTCCACCTCCCTCCCATCAG 1860
QY 2204 ccccatatataatgcccacaaatctccacaaagcctgtgctgtcctgcaaaccttactctg 2263
DB 1861 CCCCATATATATGCCCAATCTCCACAAAGCCTGCTGCAAAACCTTACTCTG 1920
QY 2264 aaatgacttccagagctggagagc 2286
DB 1921 AAATGACTTCCAGGCTGGAGC 1943

```

```

RESULT 6
AF276916 7025 bp DNA linear PRI 20-DEC-2000
LOCUS Homo sapiens melanoma differentiation association protein 7 gene,
DEFINITION complete cds.
VERSION AF276916
KEYWORDS AF276916.1 GI:11907920
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 7025)
AUTHORS Peat,J., Kude,D., Eskdale,J., Jueliger,S. and Gallagher,G.
TITLE The human MDA-7 gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7025)
AUTHORS Peat,J., Kude,D., Eskdale,J., Jueliger,S. and Gallagher,G.
JOURNAL Direct Submission
TITLE Submitted (09-JUN-2000) Department of Surgery, University of
JOURNAL Glasgow, Queen Elizabeth Building, Glasgow Royal Infirmary, Glasgow
G31 2ER, Scotland
FEATURES
source
1..7025
/organism="Homo sapiens"
/db_xref="taxon:9606"
join(<1804..1847,3285..3480,4261..4323,5460..5618,
5964..6038,6942..>7025)
/product="melanoma differentiation association protein 7"
join(1804..1847,3285..3480,4261..4323,5460..5618,
5964..6038,6942..7025)
/codon_start=1
/product="melanoma differentiation association protein 7"
/protein_id="AAG41401.1"
/db_xref="GI:11907921"
/translation="NNFOORLOSMTLARPCEPPLAPASOMQVYVPLCGFTLLMS
QVSGAGGSEPHGPGQVGVYVPOKLMFAMVXKDMOAOHITSAKLQOEVDONSD
ASCVYLHLLRFLRYKTVKYNHNTGVRLKESFSLANNFVLIYSOLOPSENEMF
SINDSNHREPLFRFRKQDYEALTALEVDVILLTNMQRKPL"
BASE COUNT 1759 a 1668 c 1759 g 1838 t 1 others
ORIGIN
Query Match 59.7%; Score 1365.2; DB 9; Length 7025;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1425; Conservative 0; Mismatches 14; Indels 22; Gaps 3;
QY 848 taccctgtccccaagtcacattcctactccglttcccaagcctggaacaggttggc 907
DB 1 TATCCTGTCCCCCGACGTCACATTCCTCCCTCCATGCTGACCCAGGTGGGC 60
QY 908 aaactctcctgttaagaacagcagaactatlttaggcctctgtgacatagtgtct 967
DB 61 AAACCTTCTCTTAAGAACACAGACAGAACTATTAGGCGCTGTGACATAGGCT 120
QY 968 cagtcacactactcactcctcctcctgtgacagaaagaaatgacaaatattgtcaa 1027
DB 121 CAGTCACAATCACTCATCTGCTGTGACAGAAACATTTGCAACATATGTCAA 180
QY 1028 caaacatagtgaaccatgaacaaacttatattatgtgatacgaagaaacctgaaataat 1087
DB 181 CAACATATGTGACCCCATGAAAACCTTATTATTATTGTGATAGGAAACCTGAAATAT 240
QY 1088 gctcttcttctgatttlttcccaatcaataaagaaactactccttaagtgcga 1147
DB 241 GCTCTTCTTTGATTTTTCCTCCATCATTTAAACAGTAACCTACTAGTGCACA 300
QY 1148 aggttaagcactcctcagcttagcagtgagcagctgagattgctgtgacactaagctg 1207
DB 301 AGGTTAAACCAATTCTCACTTAGCAGTGGAGCTGATTTGGCTTGTGACCTACAGTTG 360
QY 1208 gccaatccctgattcccaaaatgattcctcagagagatgtgggcaataactatagggaat 1267
DB 1208 gccaatccctgattcccaaaatgattcctcagagagatgtgggcaataactatagggaat 1267

```

DB 361 GCCAATCCCGATTCCCAAAATGTAATTCCTGACGATGTGGGCAAAATCTATTGGAAT 420

QY 1268 gctgattaaacagaglttaagaagcatcagacatltccagagcgagctagacatgtccag 1327

DB 421 GCTGATTTAAACAGATTAAGAACATCAAGACATTTCCAGAGAGGCGTACGACATGCCAG 480

QY 1328 ggtctctaaatgacctcat-ggattcaatcgtttcattagggaggtcttgcagaagaa 1386

DB 481 GGCTCTCACTGACCTGATAGATTCATGCTTTTCATGAGGAGATCTTGCAAGACAAGA 546

QY 1387 attcccaaacctagagctcagagagctgtgcttgggaaacactgtctcgtcttagtccc 1446

DB 541 ATTCTCAAACTAGACTGAGAGACTGAGTCTGTTGGGAAACATGCTGCTGATGCC 600

QY 1447 tcaatgggacacatgtagaaatcagagctgtgagcttgcagctagcagagagagagag 1506

DB 601 TCACGGGACATGTAGAAATCTAGAGCTGATGCTTCTGACGTGAGATAGGCTCAGA 660

QY 1507 gctcttgagcgtcctgtgacgtlcttgacacatacagctgtctgtccctgagtggtca 1566

DB 661 GCTCTTGACTGCTCTGCGACGTCTTGACATCAACGCTGTCTGTCCCTGAGTGTTC 720

QY 1567 gaggcacacagcagaagactagccacccagagacacagcctccagcttctgagcttg 1626

DB 721 GAGGCACACAGGCAAGACTAGCCACACAGACAGCAGGCTCCAGGCTTCTGCGCTTG 780

QY 1627 tccatgggtacatctctcttcttcttcttcttcttcttcttcttcttcttcttctt 1686

DB 781 TCCATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840

QY 1687 gttgagtgatataaaccttgaagacatctggtagcgtgca tgcagagaaacatccctcaac 1746

DB 841 GTTAGATGATATTAACCTAGGAGCATGGGTAGCTGATGATGATGATGATGATGATGAT 900

QY 1747 ttccggccctgagatgagtaagaagagagagagagagagagagagagagagagagag 1803

DB 901 TTCTGGGCTGATGATTAAGAGAGAGTACTTACAGGTATTTCTTCTGCTATTACGGA 960

QY 1804 ctactgtccacagca tgaatgagcatcatatgtatatttgcagagagagagagagagag 1863

DB 961 CTACTGCTCCAGCATGATGAGGATCATATTTGATTTTGTGACAGAGAGAGAGAGAG 1020

QY 1864 tgaatgagcagagtaataatgtatataatgtatataatgtatataatgtatataatgtat 1923

DB 1021 TGTATGCTGACAT 1080

QY 1924 ttgttttgccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 1985

DB 1081 TTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

QY 1966 aattatataatgag 2025

DB 1141 AATTATATCATGACAGAGACTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200

QY 2026 ggcagtggtatagagctgag 2085

DB 1201 GGCATGAGTAAAGCTGATCAACCAAGCCAGGCTTATCTACCTTACAGGCGCTTCTCC 1260

QY 2086 catcgctgactgtcctttaaagtgtatgtatgtatgtatgtatgtatgtatgtatgtatgt 2145

DB 1261 CATGCGTGAATGTGCTTAAAGAGATGAGTACTAGCAATGAGTAAATTCCTAGGAAA 1320

QY 2146 gcatgacaaatltcccttcttccactccttcttcttcttcttcttcttcttcttcttcc 2205

DB 1321 GCATGACCAATTTCCCTTTTCCACCTCCTCTTTTCTTCTTCTTCTTCTTCTTCTTCT 1380

QY 2206 cccatataatgacaaatltcccaaaagcctgtgctgtcgaacatcttactctaa 2265

DB 1381 CCGATATATATGCGCAAAATCTGCAAAAGCCTTGTGCTGCAAACTTTACTTCTGAA 1440

QY 2266 atgacttccagcgtgagag 2286

DB 1441 ATGACTTCCAGCGCTGGAGC 1461

RESULT 7  
HSA243947  
LOCUS  
DEFINITION  
Homo sapiens VCX-A gene for variably charged X chromosome protein.  
ACCESSION  
M243947.1  
VERSION  
M243947.1  
KEYWORDS  
VCX-A gene.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 (bases 1 to 29034)  
Li, X.M., Yen, P.H. and Shapiro, L.J.  
Characterization of a low copy repetitive element S232 involved in  
the generation of frequent deletions of the distal short arm of the  
human X chromosome  
Nucleic Acids Res. 20 (5), 1117-1122 (1992)  
MEDLINE  
92195814  
REFERENCE  
2 (bases 1 to 29034)  
Fukami, M., Kirsch, S., Schiller, S., Richter, A., Benes, V., Franco, B.,  
Muroya, K., Rao, E., Merker, S., Niesler, B., Ballabio, A., Ansoorge, W.,  
Ogata, T. and Rappold, G.A.  
A member of a gene family on Xp22.3, VCX-A, is deleted in patients  
with X-linked nonspecific mental retardation  
Am. J. Hum. Genet. 67 (3), 563-573 (2000)  
TITLE  
JOURNAL  
MEDLINE  
20395828  
REFERENCE  
3 (bases 1 to 29034)  
Richter, A.  
Direct Submision  
Submitted (10-AUG-1999) Richter A., Biochemical Instrumentation,  
EMBL, Meyerhofstrasse 1, Heidelberg, 69117, GERMANY  
FEATURES  
source  
1. 29034  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="Xp22.3"  
/map="between 40Bt and DXS1139"  
/clone="LINC110"  
/sex="male"  
/cell\_line="Lymphoblastoid"  
/clone\_lib="Lawrence Livermore X-chromosome specific  
library"  
/note="Variably charged, X chromosome mRNA on CRI-S232A"  
10574. 10732  
/gene="VCX-A"  
/number=1  
10574. 11947  
/gene="VCX-A"  
10733. 11046  
/gene="VCX-A"  
/number=1  
11047. 11294  
/gene="VCX-A"  
/number=2  
join(11195. 11294,11487. 11947)  
/gene="VCX-A"  
/codon\_start=1  
/product="VCX-A protein"  
/protein\_id="CAB56197.1"  
/translation="MSPKRASGAPKATENGKSSQSPSPDKTKTKYAKKGA  
VRNRRGKKGAAATKAAATYAPABESGPAAPSPQDELPOHELPEEPVSEGTQHD  
PLQSELEEPLOSQSEVPEPLQSEVPEPLQSEVPEPLQSEVPEPLQSEVPEPLQSEV  
EPLQSEVPEPLQSEVPEPLQSEVPEPLQSEVPEPLQSEVPEPLQSEVPEPLQSEV"  
11295. 11486  
/gene="VCX-A"  
/number=2  
11487. 11947  
/gene="VCX-A"  
/number=3  
11697. 11936  
repeat\_region







```

repeat_region /rpt_family="LINE2"
complement(11782..11841)
/rpt_family="AT_rich"
repeat_region complement(19230..19258)
/rpt_family="AT_rich"
repeat_region complement(19446..19742)
/rpt_family="AluY"
repeat_region 19749..19785
/rpt_family="AT_rich"
complement(20413..20518)
/rpt_family="MER5B"
repeat_region 21555..21649
/rpt_family="MER5A"
complement(21855..21892)
/rpt_family="LINE2"
repeat_region complement(22871..22908)
/rpt_family="AT_rich"
complement(23168)
/rpt_family="LINE2"
repeat_region complement(23296..23654)
/rpt_family="LINE2"
repeat_region 23894..24398
/rpt_family="MLT2CA"
complement(24536..24675)
/rpt_family="MIR"
repeat_region complement(24946..25005)
/rpt_family="MIR"
repeat_region 26012..26140
/rpt_family="LINE2"
repeat_region 26342..26477
/rpt_family="MIR"
repeat_region 26712..26929
/rpt_family="MER58A"
complement(27263..27327)
/rpt_family="MIR"
repeat_region complement(28097..28886)
/note="Sequence in this region could not be obtained from
the other strand or with an alternative sequencing
chemistry."
repeat_region 28421..28587
/rpt_family="MIR"
repeat_region 29156..29359
/rpt_family="L1"
repeat_region 30191..30324
/rpt_family="MIR"
repeat_region 30411..30728
/rpt_family="MER7A"
complement(30998..31061)
/rpt_family="LINE2"
repeat_region complement(31629..31659)
/rpt_family="AT_rich"
repeat_region 33484..34739
/rpt_family="L1"
repeat_region 34740..35075
/rpt_family="THE1B"
complement(35105..35297)
/rpt_family="MIR"
repeat_region complement(35585..35766)
/rpt_family="MIR"
repeat_region complement(36289..36589)

```

Query Match 4.08; Score 91.2; DB 9; Length 137769;  
Best Local Similarity 68.58; Pred. No. 2.7e-14;  
Matches 126; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

OY 912 tctctgttaagaacagcaagacttttaagctctgtcgcatactgctcagt 971
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44990 TTTTCCATGTAAGGCCAGATAGTAACATGTTAGCTCGGGCCATATGCTCAGT 44931
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 972 cacaactactcatctctgctctgttagcagcaagaatagcacaataatgcaaaa 1031
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44930 TGCAACTACTCACTGCTTGTGCTTGTGAGTGGGAAAGCATCACAGCAATATGCACAA 44871
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

OY 1032 catatgaccaccatgaacatttatattatgatacgaaccigaaataatgct 1091
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44870 TAGGGGATGACTGTTCCATGTAACATTTGTTACAGATGCTGAATTTATATTA 44811
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1092 tctc 1095
|||||
DB 44810 TTTT 44807
|||||

```

RESULT 10  
AL139275  
LOCUS  
DEFINITION  
Human DNA sequence from clone RP11-552E20 on chromosome  
6p12.3-21.2, complete sequence.  
ACCESSION  
AL139275  
VERSION  
AL139275.30 GI:15485119  
KEYWORDS  
FTG.  
SOURCE  
ORGANISM  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
Smith,M.  
TITLE  
Submitted (04-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
COMMENT  
Requests: clone requests@sanger.ac.uk  
On Sep 6, 2001 this sequence version replaced gi:14800199.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; Sw,  
SWISSPROT; Tr, TrEMBL; Wp, WormRep; Information on the WormRep  
database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormrep  
This sequence was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6  
RP11-552E20 is from the library RPI-11.2 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP11-552E20. The true  
left end of clone RP11-552E20 is at 161600 in this sequence. The  
true right end of clone RP11-200E19 is at 35574 in this sequence.  
FEATURES  
source  
1. 193126  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="P12.3-21.2"  
/clone="RP11-552E20"  
/clone\_id="RPI-11.2"  
12. 455  
/note="match: GSS: Em:A0807014"  
4253. 4780  
/note="match: GSS: Em:B71338"  
4256. 4678  
/note="match: GSS: Em:A0437688"  
complement(8999..9566)  
misc\_feature  
misc\_feature  
misc\_feature

```

misc_feature /note="match: STS: Em: G63453"
9829. 10237
/note="match: GSS: Em: A0030932"
misc_feature complement(10116, .10617)
/note="match: STS: Em: G55076
match: GSS: Em: A0318725"
misc_feature 13203. 13321
/note="match: GSS: Em: A0792274"
13338. 13428
/note="match: GSS: Em: A0042270"
misc_feature complement(13332, .13469)
/note="match: GSS: Em: A0504765"
13417. 13794
/note="match: GSS: Em: A0138303"
13426. 15971
/note="match: GSS: Em: A0384582"
15435. 16067
/note="match: GSS: Em: A0198235"
20235. 20758
/note="match: GSS: Em: A0480601"
20275. 20620
/note="match: GSS: Em: A0097268"
33115. 33591
/note="match: GSS: Em: A0613320"
misc_feature complement(35079, .35571)
/note="match: GSS: Em: A0692857"
misc_feature complement(38680, .39467)
/note="match: GSS: Em: A0892878"
41966. 42582
/note="match: GSS: Em: A0208209"
41976. 42381
/note="match: GSS: Em: A0124355"
46566. 46984
/note="match: GSS: Em: A0882156"
47399. 47848
/note="match: GSS: Em: A0811844"
64782. 65002
/note="match: GSS: Em: A0772609"
64783. 65002
/note="match: GSS: Em: A0683445"
/note="match: GSS: Em: A068189"
/note="match: STS: Em: G08189"
misc_feature complement(64864, .64983)
/note="match: GSS: Em: AL328221"
misc_feature complement(64864, .64975)
/note="match: GSS: Em: AL324222"
misc_feature complement(64888, .64983)
/note="match: GSS: Em: A0332064"
64898. 65015
/note="match: STS: Em: G09742"
misc_feature complement(64898, .64990)
/note="match: GSS: Em: A2257084"
misc_feature complement(64898, .64989)
/note="match: GSS: Em: A0674454"
misc_feature complement(64900, .64983)
/note="match: GSS: Em: A237938"
64903. 65013
/note="match: GSS: Em: A237938"
64903. 65013
/note="match: GSS: Em: A191898"
64903. 64991
/note="match: GSS: Em: AL309316 Em: AL316981 Em: AL346021"
64903. 64983
/note="match: GSS: Em: AL224697"
misc_feature complement(64903, .64991)
/note="match: GSS: Em: AL284370"
misc_feature complement(64903, .64983)
/note="match: STS: Em: G10109"
64905. 65002
/note="match: GSS: Em: A0284087"
misc_feature complement(74424, .74751)
/note="match: GSS: Em: A0151384"
misc_feature complement(77154, .77547)
/note="match: GSS: Em: A0192713"
80797. 81303

```

```

misc_feature /note="match: GSS: Em: A0606690"
complement(80886, .81417)
/note="match: GSS: Em: A0602297"
misc_feature complement(91202, .91756)
/note="match: GSS: Em: A0606330"
misc_feature complement(102120, .102591)
/note="match: GSS: Em: A0177699"
107123. 107721
/note="match: GSS: Em: A0524997"
complement(115935, .116552)
/note="match: GSS: Em: A0080194"
misc_feature complement(116067, .116478)
/note="match: GSS: Em: A0031004"
complement(116195, .116555)
/note="match: GSS: Em: B32716"
complement(128470, .128806)
/note="match: GSS: Em: B36077"
complement(128811, .129176)
/note="match: STS: Em: H5164XF4"
139541. 139825
/note="match: GSS: Em: A0278260"
165778. 165946
/note="match: STS: Em: G26303"
171874. 172092
/note="Single clone region. Sequence from reads from a
short insert library derived from a clone PCR.
Restriction digest data confirm the assembly."
complement(175807, .176203)
/note="match: GSS: Em: A0882642"
176202. 176743
/note="match: GSS: Em: A0374550"
complement(177911, .178307)
/note="match: GSS: Em: A0131194"
complement(177997, .178301)
/note="match: GSS: Em: A0148743"
178325. 178821
/note="match: GSS: Em: A0420001"
complement(179120, .179217)
/note="match: GSS: Em: A0545214"
BASE COUNT 54062 a 43323 c 43111 g 52630 t
ORIGIN

```

```

Query Match 4.0%; Score 90.8; DN 9; Length 193126;
Best Local Similarity 55.7%; Pred. No. 3.5e-14;
Matches 194; Conservative 0; Mismatches 152; Indels 1;
Gaps 1;

QY 901 gtcggcaaacctctccgtlaaagaccagaagaactatttagctctgtgccaat 960
DB 125379 GGTGAGCAAAATTTTCTGTAAAGAACAGATGTCATATTTTGTGGTGTGGCCAT 125438

QY 961 atggtctcgaacaaactactctctcctctgacacgaaagcaattagcaacat 1020
DB 125439 ATATCTCTGTGACCACTGCTACCTCCGCAATTCAGATGTGAAGAGCTGTGAACAA 125498

QY 1021 atgtcaacaacatattgaccccatgaataactattattatattgatacgaaactga 1080
DB 125499 ACGTAATGAACGGGCGCTGTGCCCGTAACATTTATGGCACACATGAATTTAA 125558

QY 1081 aaataagctcttctttttagtttcccaacttaaaacgttaaaactactctta 1140
DB 125559 TTTTCATCTGTGCAAAAGTATTTTCTTTGTGATTTTCCACCAATTAATAATGAAAAAC 125618

QY 1141 ggtcgaaggttaagcaatctcagcttagcagtgagcgaggtgatttgcttgacct 1200
DB 125619 CTTCCTTAACCTTGAGGGGTAGAACAAACAG--GCATCGGGCTGGATTTGGCTCATGGCC 125676

QY 1201 acagttggcaatcccgatcccaaaatgatatcttcctcaggaatgtg 1248
DB 125677 ATTGTGTGCCGACCCCTGATGTGACATCATCCCATCTCACAATGTG 125724

```

RESULT 11

LOCUS	AC025298	17957 bp	DNA	linear	HTG 03-JUN-2000
DEFINITION	Homo sapiens clone RP11-2N21, WORKING DRAFT SEQUENCE, 5 unordered pieces.				
ACCESSION	AC025298				
VERSION	AC025298.3	GI:8225254			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Barren,B., Linton,L., Nusbaum,C. and Lande,E.				
TITLE	Homo sapiens, clone RP11-2N21				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 177957)				
AUTHORS	Barren,B., Linton,L., Nusbaum,C., Lande,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barre,N., Bastien,V., Bede,F., Boguslavsky,L., Bouhagalter,B., Brown,A., Burkett,G., Campione,A., Castle,A., Choepe,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dextrallia,C., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galdag,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,U., Larocque,K., Lamazares,R., Landers,T., Lehoccky,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McKean,P., McGuck,A., McKernan,K., McSheeters,R., Meldrum,J., Meneses,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirelli,A., Travers,M., Triggilio,J., Vassilev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zoody,M.				
TITLE	Direct Submissions				
JOURNAL	Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
COMMENT	On Jun 3, 2000 this sequence version replaced gi:7670170. All repeats were identified using RepeatMasker: Smit,A.F.A. & Green,P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html				
-----					
Genome Center					
Center: Whitehead Institute/ MIT Center for Genome Research					
Center code: WIBR					
Web site: http://www-seq.wi.mit.edu					
Contact: sequence_submissions@genome.wi.mit.edu					
-----					
Project Information					
Center project name: L2680					
Center clone name: 2.N.21					
-----					
Summary Statistics					
Sequencing vector: M13; M77815; 100% of reads					
Chemistry: Dye-terminator Big Dye; 100% of reads					
Assembly program: Phrap; version 0.960731					
Consensus quality: 174723 bases at least Q40					
Consensus quality: 176598 bases at least Q30					
Consensus quality: 177197 bases at least Q20					
Insert size: 183000; agarose-fp					
Insert size: 177557; sum-of-contigs					
Quality coverage: 6.7 in Q20 bases; agarose-fp					
Quality coverage: 6.8 in Q20 bases; sum-of-contigs					
-----					
NOTE: This is a 'working draft' sequence. It currently					
* consists of 5 contigs. The true order of the pieces					
* is not known and their order in this sequence record is					
* arbitrary. Gaps between the contigs are represented as					
* runs of N, but the exact sizes of the gaps are unknown.					
* This record will be updated with the finished sequence					
* as soon as it is available and the accession number will					
* be preserved.					
1					
1828: contig of 1828 bp in length					

1 (bases 1 to 178477)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 14, clone RP11-354A24  
 unpublished  
 2 (bases 1 to 178477)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

1 (bases 1 to 178477)  
Bjirren, B., Linton, L., Nusbaum, C. and Lander, E.  
Human sapiens chromosome 14, clone RP11-354A24  
Unpublished  
2 (bases 1 to 178477)  
Bjirren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Beckery, R., Bede, F.,  
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,  
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeRellano, K., Dewar, K., Domino, M., Doyle, M., Fennestor, J.,  
Ferrelia, P., Fitzhugh, N., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, C., Klein, J.,  
Landers, T., Lehocsky, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGuck, A., McKernan, K.,  
McPheeters, R., Meldrum, J., Menais, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, A. N., Subramanian, A., Talamas, J., Tsafaye, S., Theodore, J.,  
Tjirrell, A., Vassiliadis, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 17, 2000 this sequence version replaced gi:6980319.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
Genome Center

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 354_A_24
Center Clone name: 16051
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 172070 bases at least Q40
Consensus quality: 175121 bases at least Q30
Consensus quality: 178261 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 171277; sum-of-contigs
Quality coverage: 5.4 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently
consists of 13 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence,
as soon as it is available and the accession number will
be preserved.
1      1693: contig of 1693 bp in length
1694 1793: gap of      100 bp
1794 4078: contig of 2285 bp in length
4079 4178: gap of      100 bp
4179 14243: contig of 10065 bp in length
14244 14343: gap of      100 bp
14344 21787: contig of 7444 bp in length
21788 21887: gap of      100 bp
21888 31116: contig of 9229 bp in length
31117 31216: gap of      100 bp
31217 39038: contig of 7822 bp in length
39039 39138: gap of      100 bp
39139 46809: contig of 7671 bp in length
46810 46909: gap of      100 bp
46910 57658: contig of 10649 bp in length
57659 57658: gap of      100 bp

```

	*	57659	73890:	contig of 16222 bp in length
	*	73891	73990:	gap of 100 bp
	*	73991	96635:	contig of 22665 bp in length
	*	96636	96735:	gap of 100 bp
	*	96736	120417:	contig of 23682 bp in length
	*	120418	120517:	gap of 100 bp
	*	120518	149372:	contig of 28855 bp in length
	*	149373	149472:	gap of 100 bp
	*	149473	178477:	contig of 29005 bp in length.
FEATURES				
source				
			/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
			/chromosome="14"	
			/map="14"	
			/clone="RP11-354A24"	
			/clone_1lb="RPC1-11 Human Male BAC"	
			1..1693	
misc_feature				
			/note="assembly_fragment"	
			1794..4078	
misc_feature				
			/note="assembly_fragment"	
			4179..14243	
misc_feature				
			/note="assembly_fragment"	
			14344..21787	
misc_feature				
			/note="assembly_fragment"	
			clone_end:r7	
			vector_side:right"	
misc_feature				
			21888..31116	
misc_feature				
			/note="assembly_fragment"	
			31217..39038	
misc_feature				
			/note="assembly_fragment"	
			39139..46809	
misc_feature				
			/note="assembly_fragment"	
			clone_end:sp6	
			vector_side:right"	
misc_feature				
			46910..57558	
misc_feature				
			/note="assembly_fragment"	
			57659..73890	
misc_feature				
			/note="assembly_fragment"	
			73991..96635	
misc_feature				
			/note="assembly_fragment"	
			96736..120417	
misc_feature				
			/note="assembly_fragment"	
			120518..149372	
misc_feature				
			/note="assembly_fragment"	
			149473..178477	
misc_feature				
			/note="assembly_fragment"	
BASE COUNT		51444	a	37184
ORIGIN			c	36608
			g	52041
			t	1200
			others	
Query Match			3.9%;	Score 90; DB 2; Length 178477;
Best Local Similarity			61.5%;	Pred. No. 5.9e-14;
Matches 182; Conservative			0;	Mismatches 105; Indels 9; Gaps 2;
OY	861	acgtcacattcctactcccglttcccagtcgcgtgaaccaggctgggcaaaccttccctgt	920	
Db	10108	AGGTTTATTGCTAGTCGTGGCAAAAGAACCAAGACAAGAAAGCGTCACTGAATCTTTCTGT	10049	
OY	921	aaagaaccagaagaactatcttgagtcgtgtgcatatgctctcagccaactac	980	
Db	10048	GAAAGACTAGTAATTAATTTTGAGSGTTTGACGACCATATGTGCTCTTGCAACTAC	9989	
OY	981	tcaatctgcctctgtagcagaaagcaaatagacaataatgtcacaacaatatgtga	1040	
Db	9988	TCAAAGTCTCTCTCTAGTAGTAAGAAGCCACCGTGAGCAATAATGTMAATGAATGGGCATGA	9929	
OY	1041	ccccat--gaaaacttatattatctgatgacggaaccggaataatagtctttc----	1094	
Db	9928	CTGCATTGCCAATTAATTTTCTTTATGACATGACATTGACATTTTATACATGTCAGA	9869	
OY	1095	-----tttgtatcttttccccaatcataaaaaacgtaaaaacctcttggtgcga	1147	



```

VERSION      AC011060.5  GI:7229784
KEYWORDS     HTGS_PHASE1; HTGS_DRAFT.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE        1 (bases 1 to 181009)
JOURNAL      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE    Homo sapiens, clone RP11-11G16
AUTHORS      2 (bases 1 to 181009)
JOURNAL      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
REFERENCE    Blotwain,J., Barua,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
AUTHORS      Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
REFERENCE    Cooke,P., Dearliano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
AUTHORS      Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D.,
REFERENCE    Galagan,J., Gardina,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
AUTHORS      Howland,J.C., Johnson,R., Jones,C., Kang,L., Karatas,A., Klein,J.,
REFERENCE    Lebecky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
AUTHORS      McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
REFERENCE    Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
AUTHORS      Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
REFERENCE    Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
AUTHORS      Tsefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
REFERENCE    Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome
COMMENT      Research, 330 Charles Street, Cambridge, MA 02141, USA.
              On Mar 12, 2000 this sequence version replaced g1:6964405.
              All repeats were identified using RepeatMasker:
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              Smith, A.F.A. & Green, P. (1996-1997)
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: L3156
              Center clone name: 11.G.16
              ----- Summary Statistics
              Sequencing vector: M13; M77815; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 173981 bases at least Q40
              Consensus quality: 172073 bases at least Q30
              Consensus quality: 175479 bases at least Q20
              Insert size: 155000; agarose-fp
              Insert size: 178609; sum-of-ctrls
              Quality coverage: 5.1 in Q20 bases; agarose-fp
              Quality coverage: 4.5 in Q20 bases; sum-of-ctrls
              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 25 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              1
              * 1503 1604: contig of 1504 bp in length
              * 1505 1604: gap of 100 bp
              * 1605 2791: contig of 1187 bp in length
              * 2792 2891: gap of 100 bp
              * 2892 3909: contig of 1018 bp in length
              * 3910 4009: gap of 100 bp
              * 4010 5556: contig of 1547 bp in length
              * 5557 5656: gap of 100 bp
              * 5657 7482: contig of 1826 bp in length
              * 7483 7582: gap of 100 bp
              * 7583 8926: contig of 1344 bp in length
              * 8927 9026: gap of 100 bp
              * 9027 10470: contig of 1444 bp in length

```

```

FEATURES
source
* 10471 10570: gap of 100 bp
* 10571 12198: contig of 1629 bp in length
* 12200 12299: gap of 100 bp
* 12300 14087: contig of 1788 bp in length
* 14088 14187: gap of 100 bp
* 14188 15707: contig of 1520 bp in length
* 15708 15807: gap of 100 bp
* 15808 17945: contig of 2138 bp in length
* 17946 18045: gap of 100 bp
* 18046 20017: contig of 1972 bp in length
* 20018 20117: gap of 100 bp
* 20118 22115: contig of 1998 bp in length
* 22116 22215: gap of 100 bp
* 22216 24556: contig of 2341 bp in length
* 24557 24656: gap of 100 bp
* 24657 27568: contig of 2912 bp in length
* 27569 27668: gap of 100 bp
* 27669 32527: contig of 4859 bp in length
* 32528 32627: gap of 100 bp
* 32628 37892: contig of 5265 bp in length
* 37893 37992: gap of 100 bp
* 37993 41265: contig of 3273 bp in length
* 41266 41365: gap of 100 bp
* 41366 52400: contig of 11035 bp in length
* 52401 52500: gap of 100 bp
* 52501 63506: contig of 11006 bp in length
* 63507 63606: gap of 100 bp
* 63607 79362: contig of 15756 bp in length
* 79363 79462: gap of 100 bp
* 79463 97023: contig of 17561 bp in length
* 97024 97123: gap of 100 bp
* 97124 124426: contig of 27303 bp in length
* 124427 124526: gap of 100 bp
* 124527 149337: contig of 24811 bp in length
* 149338 149437: gap of 100 bp
* 149438 181009: contig of 31572 bp in length.
              Location/Qualifiers
              1. 181009
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="RP11-11G16"
              /clone_id="RP11-11G16"
              1. 1504
              /note="assembly-fragment"
              1605. 2791
              /note="assembly-fragment"
              2892. 3909
              /note="assembly-fragment"
              4010. 5556
              /note="assembly-fragment"
              5557. 7482
              /note="assembly-fragment"
              7583. 8926
              /note="assembly-fragment"
              9027. 10470
              /note="assembly-fragment"
              10571. 12199
              /note="assembly-fragment"
              12300. 14087
              /note="assembly-fragment"
              14188. 15707
              /note="assembly-fragment"
              15808. 17945
              /note="assembly-fragment"
              18046. 20017
              /note="assembly-fragment"
              20118. 22115
              /note="assembly-fragment"
              22216. 24556
              /note="assembly-fragment"
              24657. 27568
              /note="assembly-fragment"
              27669. 32527
              /note="assembly-fragment"

```



```
misc_feature /note="assembly_fragment"  
32628 .37892  
/note="assembly_fragment"  
misc_feature 37993 .41265  
/note="assembly_fragment"  
clone_end:SP6  
vector_side:right  
misc_feature 41366 .52400  
/note="assembly_fragment"  
misc_feature 52501 .63506  
/note="assembly_fragment"  
misc_feature 63607 .79362  
/note="assembly_fragment"  
clone_end:T7  
vector_side:right  
misc_feature 79463 .97023  
/note="assembly_fragment"  
misc_feature 97124 .124426  
/note="assembly_fragment"  
misc_feature 124527 .149337  
/note="assembly_fragment"  
misc_feature 149438 .181009  
/note="assembly_fragment"
```

BASE COUNT    57610 a   36532 c   33173 g   51292 t       2402 others

ORIGIN

Query Match	3.98	Score 89.2	DB 2	Length 181009
Best Local Similarity	64.68	Pred. No. 9.9e14		
Matches 133; Conservative	0	Mismatches 73;	Indels 0	Gaps 0

QY 891 ctgaccacaggttggcaactcttcctgtaagaaccagacgaactatttagctc 950  
||| | | |||||| | |||||| | |||||| |  
Db 151319 CTAGACACAGGCTCAGCCAACTTCTCTGTATAAGACCAGATCATTAATATTTTAGGTTT 151378

QY 951 tgtgtgccatagtctcagcacaactactatctctgcctctgttagcagaaagcaat 1010  
||||| ||||||| ||||||| ||||||| ||||| ||||| ||||||| ||  
Db 151379 TGTGGGCCATATGCTCTCTGTCAACAACACTACTCTCACTCTGCTGTTCTTAGGGCAAAAGCAGT 151438

QY 1011 tagcaacataatgtcacaacataatgtgaccocatgaaccttatattatgatcac 1070  
||| ||||| | || || | | | ||||| ||  
Db 151439 TATAGAAAATACATAAATGAATGAGCATGGCTGTGCCATAAAGATTATTATGACAC 151498

QY 1071 ggaacctgaaataatgtcttctt 1096  
 |||| |||| | ||| | ||  
 Db 151499 TGAATTTGAAATTCATGTAATTTT 151524

RESULT 15  
A1161628

LOCUS	163453 bp	DNA	11kb	TRI 10-CON 200
AL161628				
DEFINITION				
Human DNA sequence from clone Rp11-3K16 on chromosome 9. Contains a snRNA binding domain pseudogene, the EAVL2 gene for E1AV (embryonic lethal, abnormal vision, Drosophila)-like 2, STSS				

ACCESSION  
GSSS and a CPG island, complete sequence.  
AL161628  
VERSION  
AL161628.9 GI:10129841  
KEYWORDS  
HMG; CPG island; FLAVR2; snORNA binding domain.

SOURCE	ORGANISM	
human.	human.	1
Homo sapiens	Chordata: Vertebrata: Euteleostomi;	
Eukaryota; Metazoa:	Carnifera: Carnifera: Carnifera: Carnifera:	
Mammalia: Eutheria:	Primates: Carnifera: Carnifera: Carnifera:	
	Homidae: Homidae: Homidae: Homidae:	

REFERENCE 1 (bases 1 to 183483)  
AUTHORS Ramsay H.  
TITLE Direct Submission  
JOURNAL submitted (20-nov-2000)  
INSTITUTION Sanger Centre  
LOCATION Hinxton, Cambridgeshire, UK

COMMENT  
On Sep 14, 2000 this sequence version replaced g1:10039694.  
Differences assembly data is compared from overalpinning clones  
requests: clonerequests@anger.ac.uk  
CB10 1st, UK. E-mail enquiries: humphreys@anger.ac.uk Clonerequests: clonerequests@anger.ac.uk

any differences are found these are annotated as variations where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. Associate primary accession numbers given in the feature table with their source databases: Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/BGP/Chr9>

This sequence is the entire insert of clone RP11-31K16 The true left end of clone RP11-32L12 is at 139671 in this sequence. The true right end of clone RP11-31S14 is at 73045 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-31K16 is from the library RPCR11.1 constructed by the group of Plierer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES	Location/Qualifiers
SOURCE	1. .183483 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="9" /clone="RP11-31K16" /clone_11b="RPC1-11.1"
misc_feature	1. .224 /note="match: GSS: Em:A0040596"
misc_feature	1. .456 /note="match: GSS: Em:A0080076"
misc_feature	24. .653 /note="match: GSS: Em:A0045921"
repeat_region	415. .471 /note="AluYo repeat: matches 60. .118 of consensus"
repeat_region	472. .763 /note="AluY repeat: matches 10. .288 of consensus"
repeat_region	764. .954 /note="AluYo repeat: matches 118. .297 of consensus"
repeat_region	1096. .1173 /note="MLTJ1 repeat: matches 103. .187 of consensus"
repeat_region	2555. 4081 /note="L1MC2 repeat: matches 4721. .6314 of consensus"
repeat_region	4184. .4223 /note="MLTJ1 repeat: matches 238. .273 of consensus"
repeat_region	4424. .4479 /note="MLTJ1 repeat: matches 343. .398 of consensus"
repeat_region	5387. 5466 /note="MER58 repeat: matches 2214. .2294 of consensus"
repeat_region	5639. .5786 /note="L1ME3A repeat: matches 6017. .6156 of consensus"
repeat_region	6111. .6324 /note="MER20 repeat: matches 1. .218 of consensus"
repeat_region	6330. .6457 /note="MER5A repeat: matches 2. .120 of consensus"
repeat_region	6840. .7315 /note="L2 repeat: matches 2252. .2736 of consensus"
repeat_region	7345. .7434 /note="AluX repeat: matches 2. .89 of consensus"
repeat_region	7437. .7741 /note="AluX repeat: matches 1. .299 of consensus"
repeat_region	7808. .7989 /note="MER58 repeat: matches 28. .2412 of consensus"
repeat_region	10366. .10649 /note="MER58B repeat: matches 46. .341 of consensus"
repeat_region	11076. .11221 /note="MLT2 repeat: matches 1. .147 of consensus"
repeat_region	11222. .11661 /note="MLT2CB repeat: matches 1. .504 of consensus"
repeat_region	11664. .11992

```

repeat_region /note="MLT2 repeat: matches 145. .510 of consensus"
12210. .12467
misc_feature /note="12 repeat: matches 2482. .2750 of consensus"
complement(12686. .13167)
misc_feature /note="match: GSS: Em:AQ549073"
complement(13792. .14266)
misc_feature /note="match: GSS: Em:AQ594257"
complement(14438. .14735)
repeat_region /note="match: GSS: Em:AQ128238"
14756. .14785
repeat_region /note="15 copies 2 mer aa 100% conserved"
16212. .16363
repeat_region /note="FRAM repeat: matches 4. .145 of consensus"
16364. .16385
misc_feature /note="11 copies 2 mer tt 100% conserved"
16462. .16951
repeat_region /note="match: GSS: Em:AQ631149"
16908. .17128
misc_feature /note="MIR repeat: matches 34. .261 of consensus"
complement(17116. .17552)
misc_feature /note="match: GSS: Em:AQ133240"
17546. .18050
misc_feature /note="match: GSS: Em:AQ777145"
17660. .17960
repeat_region /note="match: GSS: Em:AQ094704"
19777. .19993
misc_feature /note="MER58A repeat: matches 2. .224 of consensus"
20817. .21173
misc_feature /note="match: GSS: Em:AQ376698"
complement(20830. .20939)
misc_feature /note="match: GSS: Em:AQ229933"
complement(20830. .21133)
misc_feature /note="match: GSS: Em:AQ235822"
complement(20830. .21137)
repeat_region /note="match: GSS: Em:B65379"
21100. .21154
repeat_region /note="11 copies 5 mer aaac 78% conserved"
21183. .21410
repeat_region /note="MER30 repeat: matches 1. .230 of consensus"
21647. .21773
misc_feature /note="MIR repeat: matches 119. .246 of consensus"
complement(23139. .23500)
repeat_region /note="match: GSS: Em:AQ210496"
23363. .23601
misc_feature /note="MIR repeat: matches 2. .256 of consensus"
complement(24372. .24810)
repeat_region /note="match: GSS: Em:AQ178605"
26871. .27202
repeat_region /note="MSTD repeat: matches 1. .341 of consensus"
27203. .27695
repeat_region /note="LIPAL10 repeat: matches 5642. .6165 of consensus"
27696. .27750
repeat_region /note="MSTD repeat: matches 341. .394 of consensus"
27867. .27971
repeat_region /note="MERSA repeat: matches 84. .189 of consensus"
27986. .28150
misc_feature /note="MERSA repeat: matches 3. .183 of consensus"
28765. .29329
misc_feature /note="match: GSS: Em:B93338"
28777. .29151
repeat_region /note="match: GSS: Em:AQ125799"
29425. .29482
repeat_region /note="29 copies 2 mer ac 87% conserved"
29740. .29805
repeat_region /note="33 copies 2 mer aa 71% conserved"
29763. .29807
repeat_region /note="9 copies 5 mer aaac 82% conserved"
30205. .30412
repeat_region /note="MIR repeat: matches 2. .230 of consensus"
30382. .30732
misc_feature /note="MERSA repeat: matches 8. .187 of consensus"
30911. .31378
misc_feature /note="match: GSS: Em:AQ546895"

```

```

repeat_region 32009. .32080
gene /note="MIR repeat: matches 132. .206 of consensus"
32125. .32784
CDS /gene="ba31k16.1"
32125. .32784
/gene="ba31k16.1"
match: proteins: Sw:000567 Tr:Q9YMW69 Tr:Q9ZRW0 Tr:Q9YD11
Tr:094514 Sw:021276 Tr:Q9SGT7 Tr:Q9Y158 Sw:012460
Tr:080401 Tr:065334 Tr:065335 Tr:Q9Y5W4
/codon_start=1
/pseudo
/evidence=not experimental
repeat_region 34666. .34977
repeat_region /note="Alusg repeat: matches 11. .297 of consensus"
36947. .36978
misc_feature /note="16 copies 2 mer aa 87% conserved"
38608. .39160
polyA_site /note="match: GSS: Em:AQ353745"
complement(40339)
mRNA /gene="ELAVL2"
complement(join(40339. .43119,43683. .43721,51615. .51840,
55153. .55306,81257. .81360,112241. .112481,171869. .171992))
/gene="ELAVL2"
/note="match: cDNAs: Em:U29088 Em:X85111 Em:M62843
Em:U17602 Em:U12431 Em:L26405 Em:U17599 Em:U29148
Em:U17597 Em:D26158 Em:SB3320 Em:D31953 Em:AE176675
match: ESTs: Em:AA545382 Em:AA081584 Em:A1843112
Em:AA656788 Em:AA692890 Em:A1145457 Em:AV117363
Em:AW524152 Em:AA416369 Em:A1028574 Em:AM491988
Em:AA549705 Em:D61361 Em:AV362071 Em:A1837658 Em:AV094529

```

```

Query Match 3.9% Score 89.2 DB 9 Length 183483;
Best Local Similarity 64.6% Pred No. 9.9e-14;
Matches 133; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 891 ctgaaccaggttgycacaactctctctgtaaagacagacaggaactatttagctc 950
DB 10360 CTACACACAGAGCTCAGCAAACTTCTCTGTAAAGCAGCATGATTAATTTTACGTTT 10419
QY 951 tgtgtgcattatggtctcagcacaactactcaatctctgtctgttagcagaagaat 1010
DB 10420 TGTGGCCATATGTCCTCTGTGCACAACTACTCCACTCTGTCTGTGTGGGCAAAAGCAGT 10479
QY 1011 tagcaacaatagcacaacaacatagtgaccccatgaaaacttatattatattgattac 1070
DB 10480 TATGAAATAATCATTAATTAATGATGCGATGCGTGTGTTCCCATAAAGTATTATGACAC 10539
QY 1071 ggaacctgaataataatgcttctt 1096
DB 10540 TGAATTTTGAATTCATGTAATTTT 10565

```

Search completed: August 16, 2002, 06:01:15  
 Job time: 19139 sec

Mon Aug 19 06:46:31 2002

us-09-515-369b-1.rge

Page 23



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:50:31 ; Search time 97.23 seconds  
(without alignments)  
5775.154 Million cell updates/sec

Title: US-09-515-369B-1

Perfect score: 2286  
Sequence: 1 taatacgaactactactataggg.....tgacttcacagcgctggagacg 2286

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued\_Patents\_NA:\*  
2: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/Backlist1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Query Length	ID	Description
1	71.8	3.1	7218	1 US-08-232-463-14	Sequence 14, Appl
2	67.6	3.0	6235	4 US-09-305-384-5	Sequence 5, Appl
3	67.6	3.0	6679	4 US-09-305-384-1	Sequence 1, Appl
4	62.2	2.7	178	4 US-08-991-789A-210	Sequence 210, App
5	62.2	2.7	178	4 US-09-062-451-210	Sequence 210, App
6	52.6	2.3	7218	1 US-08-232-463-14	Sequence 14, Appl
7	52	2.3	3786	4 US-08-975-762-42	Sequence 42, Appl
8	52	2.3	3786	4 US-09-295-028-42	Sequence 42, Appl
9	52	2.3	3786	4 US-09-106-582-42	Sequence 42, Appl
10	50.6	2.2	502	2 US-08-592-541-91	Sequence 91, Appl
11	50.6	2.2	502	2 US-08-592-541-91	Sequence 91, Appl
12	50.6	2.2	502	3 US-09-124-698-91	Sequence 91, Appl
13	50.6	2.2	502	4 US-09-127-480-91	Sequence 91, Appl
14	50.6	2.2	502	4 US-08-496-841C-91	Sequence 91, Appl
15	50.6	2.2	1883	2 US-08-967-101-10	Sequence 10, Appl
16	50.6	2.2	1883	2 US-08-967-101-10	Sequence 10, Appl
17	50.6	2.2	1883	2 US-08-592-541-10	Sequence 10, Appl
18	50.6	2.2	1883	2 US-08-592-541-10	Sequence 10, Appl
19	50.6	2.2	1883	3 US-08-888-077A-8	Sequence 8, Appl
20	50.6	2.2	1883	3 US-08-124-698-10	Sequence 10, Appl
21	50.6	2.2	1883	3 US-09-124-698-10	Sequence 10, Appl
22	50.6	2.2	1883	4 US-09-127-480-10	Sequence 10, Appl
23	50.6	2.2	1883	4 US-09-127-480-10	Sequence 10, Appl
24	50.6	2.2	1883	4 US-08-496-841C-10	Sequence 10, Appl
25	50.6	2.2	1883	4 US-08-496-841C-10	Sequence 10, Appl
26	48.2	2.1	80246	4 US-09-078-294-4	Sequence 4, Appl
27	48.2	2.1	80595	4 US-09-078-294-4	Sequence 4, Appl

28	45.8	2.0	1734	2 US-08-858-052-2	Sequence 2, Appl
29	45.8	2.0	1734	3 US-09-200-284-2	Sequence 2, Appl
30	43.6	1.9	460	2 US-08-967-101-48	Sequence 48, Appl
31	43.6	1.9	460	2 US-08-592-541-48	Sequence 48, Appl
32	43.6	1.9	460	3 US-09-124-698-48	Sequence 48, Appl
33	43.6	1.9	460	4 US-08-496-841C-48	Sequence 48, Appl
34	43.6	1.9	460	4 US-08-496-841C-48	Sequence 48, Appl
35	42.8	1.9	59065	4 US-09-813-817-3	Sequence 3, Appl
36	42.4	1.9	289	4 US-09-007-005-17	Sequence 17, Appl
37	42.4	1.9	289	4 US-09-244-796-17	Sequence 17, Appl
38	41.4	1.8	703	4 US-09-313-300-6	Sequence 6, Appl
39	41.4	1.8	176373	3 US-09-128-155-17	Sequence 17, Appl
40	41.2	1.8	7505	4 US-09-078-294-13	Sequence 13, Appl
41	39.8	1.7	350	2 US-08-332-766A-20	Sequence 20, Appl
42	39.2	1.7	84495	4 US-09-797-906-3	Sequence 3, Appl
43	39.2	1.7	152331	3 US-09-128-155-16	Sequence 16, Appl
44	38.6	1.7	43795	3 US-08-742-185-101	Sequence 101, App
45	38.2	1.7	1732	1 US-08-480-346-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-F1s  
US-08-232-463-14

Query Match	3.18;	Score 71.8;	DB 1;	Length 7218;
Best Local Similarity	4.08;	Pred. No. 9.7e-12;		
Matches	16;	Conservative 237;	Mismatches 144;	Indels 0;
			Gaps	0;

[illegible]

```

RESULT      2
US-09-305-384-5
: Sequence 5, Application US/09305384
: Patient No. 6242218
: GENERAL INFORMATION:
:   APPLICANT: Treco, Douglas A.
:   APPLICANT: Hearlstein, Michael W.
:   APPLICANT: Seiden, Richard F.
:   TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
:   FILE REFERENCE: 07236/017001
:   CURRENT APPLICATION NUMBER: US/09/305,384
:   CURRENT FILING DATE: 1999-05-05
:   EARLIER APPLICATION NUMBER: US 60/084,649
:   EARLIER FILING DATE: 1998-05-07
:   NUMBER OF SEQ ID NOS: 8
:   SOFTWARE: FastSeq for Windows Version 3.0
:   SEQ ID NO 5
:   LENGTH: 6235
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:   US-09-305-384-5

```

Query Match	3.08;	Score 67.6;	DB 4;	Length 6235;
Best Local Similarity	63.7%;	Pred. No. 1.8e-10;		
Matches 135; Conservative	0;	Mismatches 74;	Indels 3;	Gaps 2;

Oy 1101 tttttcccaatcattaaanaacgtaanaac 1132  
 Db 741 gacatcgatcttgattcatatgctt--ttacattttaataataacttttaaa 798  
 Oy 1041 ccccaagaaacttatattaatgatacggaacctgaaatatabgtctttcttga 1100  
 Db 681 tctgcgcgtcttctctgcgaagaacgagcttaaacatacatatgaattttta 740  
 Oy 981 tcatctctgcctctgtagcagaagaacttagcacataltgcaacaacatltga 1040  
 Db 622 aagaatccagatagtaacttttttaggtc-tgtagccgatagtgtctctgtcaacatc 680  
 Oy 921 aagaacccagagaactatttttaggtcctctgtgcacatagtctcagtaaacatc 980  
 Db 622 aagaatccagatagtaacttttttaggtc-tgtagccgatagtgtctctgtcaacatc 680

Db 799 atttccctaaccatttaaaagtgtaaaagc 830

```

RESULT      3
US-09-305-384-1
? Sequence 1, Application US/09305384
? Patent No. 6242218
? GENERAL INFORMATION:
? APPLICANT: Treco, Douglas A.
? APPLICANT: Heartlein, Michael W.
? APPLICANT: Seldén, Richard P
? TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
? FILE REFERENCE: 07236/017001
? CURRENT APPLICATION NUMBER: US/09/305,384
? CURRENT FILING DATE: 1999-05-05
? EARLIER APPLICATION NUMBER: US 60/084,649
? EARLIER FILING DATE: 1998-05-07
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 1
? LENGTH: 6679
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-305-384-1

```

Query Match	3.08;	Score 67.6;	DB 4;	Length 6679;
Best Local Similarity	63.78;	Pred. No. 1.9e-10;		
Matches 135;	Conservative	0;	Mismatches 74;	Indels 3;
				Gaps 2

QY	921	aaagacacgcagaaacactttttaggcctctgtgcacatttgcttcagtcacacactc	980
Db	641	aaagatccagatgatcattttttttaggc-tgtgtgagccgatgtgtccctgtccacatcac	699
QY	981	tcactctgtccctctgtatgcagaaagcaattgcacacatactgtccaaacacatactgtga	1040
Db	700	tcgtgcctgtctctctatgcacaaagcagctataacatacatcaatgatttttttata	759
QY	1041	ccccatgtaaaacttattcttatctgtatcgcgaaacctgtaaaataatgtctttctttga	1100
Db	760	gacatcagatttgaatttcataatgtct--cttaacatttataaataaacatttttanaa	817
QY	1101	ttttttccccaatcatttaanaacgttaanaac	1132
Db	818	attttccccaatcatttaanaagtgtaanaagc	849

RESULT 4  
US-08-991-789A-210  
Sequence 210, Application US/08991789A  
Patent No. 6225054  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Filth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,789A

FILED DATE: 11-Dec-1997  
CLASSIFICATION: <UNKNOWN>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 210:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 210:  
US-08-991-789A-210

Query Match 2.7%; Score 62.2; DB 4; Length 178;  
Best Local Similarity 71.3%; Pred. No. 9,6e-10;  
Matches 82; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 891 ctgagccaggttgggaacctcttcctgaaagacagaggaactatttagctc 950  
DB 20 CTAGAGTAGGGTTGGCCACTTTTCTATTAAGGACGAGAGTAATATTTCAGGCTT 79  
OY 951 tgtgtccatattgtctcagtcacactactatctctgcctctgtacagaaa 1005  
DB 80 TGTGGGTGTGCACTCTCTCTTGCACACTACCTGCTGCCATTTGTACATAGAA 134

RESULT 5  
US-09-062-451-210  
Sequence 210, Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
APPLICANT: Smith, John M.  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-Apr-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 210:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-210

Query Match 2.7%; Score 62.2; DB 4; Length 178;  
Best Local Similarity 71.3%; Pred. No. 9,6e-10;  
Matches 82; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 891 ctgagccaggttgggaacctcttcctgaaagacagaggaactatttagctc 950  
DB 20 CTAGAGTAGGGTTGGCCACTTTTCTATTAAGGACGAGAGTAATATTTCAGGCTT 79  
OY 951 tgtgtccatattgtctcagtcacactactatctctgcctctgtacagaaa 1005  
DB 80 TGTGGGTGTGCACTCTCTCTTGCACACTACCTGCTGCCATTTGTACATAGAA 134

RESULT 6  
US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300,6  
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-F15  
US-08-232-463-14

Query Match 2.3%; Score 52.6; DB 1; Length 7218;  
Best Local Similarity 9.0%; Pred. No. 9,4e-06;  
Matches 28; Conservative 162; Mismatches 121; Indels 0; Gaps 0;

OY 484 catcacgcccttgaacattatctcatagtagttagtgggtctcttgtaagc 543  
DB 1490 CATCACTGTAAATTAATCTATCTATGCAAGTAGTTAAAGAGATGAATTTGTACRRRR 1431

	Query Match	2.3%	Score 52	DB 4	Length 3786
	Best Local Similarity	67.2%	Pred. No. 9.7e-06		
	Matches	30	Conservative	0	Mismatches 40
					Indels 4
					Gaps 1
QY	914	ttctctgaagaagcaacagagaaacattttgtctgtctgtgcataatgctc-----tca	969		
Db	3296	TTTCTGTGTACTGCCACACAGTAAACATTITTTGAACTTTTGATGTATTAATGGTCTCTTTCT	3555		
QY	970	gtccacaactatcattcctctgtctctgtltagacagaaagcaatttagcaacaatattcaca	1029		

RESULT 9  
US-09-106-582-42  
: Sequence 42, Application US/09106582  
: Patent No. 6306402  
: GENERAL INFORMATION:  
: APPLICANT: Reed, Steven G.  
: APPLICANT: Lodes, Michael J.  
: APPLICANT: Houghdon, Raymond  
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND  
: NUMBER OF SEQUENCES: 73  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SEED AND BERRY LLP  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: City: Seattle  
: STATE: Washington  
: COUNTRY: USA  
: ZIP: 98104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/106,582  
: FILING DATE: 29-JUN-1998







```

; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
;
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-08-496-841C-91

Query Match          2.2%; Score 50.6; DB 4; Length 502;
Best Local Similarity 64.3%; Pred. No. 7.6e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgtaagaaccagacaggaactatttaggcctctgtgcccataatggtcagtcacaa 976
    ||||||| ||||| | ||||||| ||||| ||||||| |||||
DB 264 CTGTAAGAACGCCAGACTG--AATATTTTAAGCTCTATGGGTCAATGCTCCAGGGCAA 321

QY 977 ctactcatctcgtcctctgtagacgaaagcaattagcaacaataatgccaacaacatat 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 322 ACACTCAACTGCTGCTACTGTGTAAGTGAAGCAGGACAGACATATGTTAACCAAGAGAG 381

QY 1037 gtgaaccatgaaaactatttt 1059
    ||| | | | | | | | | | | | | | | | | | | | | |
DB 382 GTGCTCACTTTCCATGAAAGTT 404

RESULT 15
US-08-967-101-10/c
; Sequence 10, Application US/08967101
; Patent No. 5840540
;
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
;
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1883 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-08-967-101-10

Query Match          2.2%; Score 50.6; DB 2; Length 1883;
Best Local Similarity 64.3%; Pred. No. 1.7e-05;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgtaagaaccagacaggaactatttaggcctctgtgcccataatggtcagtcacaa 976
    ||||||| ||||| | ||||||| ||||| ||||||| ||||| |||||
DB 239 CTGTAAGAACGCCAGACTG--AATATTTTAAGCTCTATGGGTCAATGCTCCAGGGCAA 182

QY 977 ctactcatctcgtcctctgtagacgaaagcaattagcaacaataatgccaacaacatat 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 ACACTCAACTGCTGCTACTGTGTAAGTGAAGCAGGACAGACATATGTTAACCAAGAGAG 122

QY 1037 gtgaaccatgaaaactatttt 1059
    ||| | | | | | | | | | | | | | | | | | | | | |
DB 121 GTGCTCACTTTCCATGAAAGTT 99
```

Search completed: August 16, 2002, 05:14:32  
Job time: 15841 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 05:55:47 ; Search time 1893.35 Seconds  
(without alignments)  
3109.111 Million cell updates/sec

Title: US-09-515-369b-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatcagctactataggg.....aatctccacaagccttgc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1664314 seqs, 1307041093 residues

Total number of hits satisfying chosen parameters: 3328628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	82.8	3.7	US-10-027-632-119945	Sequence 119945,
2	82.8	3.7	US-10-027-632-119946	Sequence 119946,
3	81.8	3.7	US-10-105-299-12107	Sequence 12107, A
4	81.4	3.6	US-10-105-299-12108	Sequence 12108, A
5	81	3.6	US-09-629-469A-12215	Sequence 12215, A
6	79.4	3.5	US-09-918-995-12893	Sequence 12893, A
7	79.4	3.5	US-09-785-276A-53453	Sequence 55453, A
8	79	3.5	US-10-027-632-151080	Sequence 151080,
9	79	3.5	US-10-027-632-151081	Sequence 151081,
10	78.2	3.5	US-10-027-632-7748	Sequence 7748, AP
11	77.2	3.4	US-10-104-047-364	Sequence 364, App
12	76.6	3.4	US-10-027-632-186700	Sequence 186700,
13	76.6	3.4	US-10-027-632-142654	Sequence 142654, A
14	76.6	3.4	US-09-785-276A-55307	Sequence 55307, A
15	75.6	3.4	US-10-027-632-233742	Sequence 233742,
16	75.6	3.4	US-10-027-632-233743	Sequence 233743,
17	75.6	3.4	US-10-027-632-149395	Sequence 149395,
18	75.6	3.4	US-10-027-632-149396	Sequence 149396,
19	74.6	3.3	US-10-027-632-279424	Sequence 279424,
20	74.4	3.3	US-10-027-632-117955	Sequence 117955,
21	73.4	3.3	US-10-027-632-279426	Sequence 279426,
22	73.4	3.3	US-10-027-632-279422	Sequence 279422,
23	73.4	3.3	US-10-027-632-214662	Sequence 214662,
24	73.2	3.3	US-10-027-632-252726	Sequence 252726,
25	73	3.3	US-10-027-632-279423	Sequence 279423,

26	72.8	3.2	428	7	US-10-027-632-301493	Sequence 301493,
27	72.8	3.2	428	7	US-10-027-632-301494	Sequence 301494,
28	72.8	3.2	428	7	US-10-027-632-301495	Sequence 301495,
29	72.8	3.2	428	7	US-10-027-632-301496	Sequence 301496,
30	72.8	3.2	445	7	US-10-027-632-84769	Sequence 84769, A
31	72.8	3.2	445	7	US-10-027-632-84770	Sequence 84770, A
32	72.8	3.2	445	7	US-10-027-632-84771	Sequence 84771, A
33	72.8	3.2	445	7	US-10-027-632-84772	Sequence 84772, A
34	72.8	3.2	791	7	US-10-027-632-214334	Sequence 214334,
35	72.2	3.2	634	7	US-10-027-632-47300	Sequence 47300, A
36	72	3.2	591	7	US-10-027-632-47301	Sequence 47301, A
37	72	3.2	608	7	US-10-027-632-114163	Sequence 114163, A
38	71.6	3.2	837	7	US-10-027-632-29465	Sequence 29465, A
39	71.6	3.2	837	7	US-10-027-632-29466	Sequence 29466, A
40	71.4	3.2	585	7	US-10-027-632-115173	Sequence 115173,
41	71.4	3.2	1327	5	US-09-919-002-3556	Sequence 3556, Ap
42	71.2	3.2	105413	8	US-60-377-592-3	Sequence 3, Appl1
43	70.8	3.2	470	5	US-09-721-544-18515	Sequence 18515, A
44	70.8	3.2	631	7	US-10-027-632-143593	Sequence 143593,
45	70.8	3.2				

ALIGNMENTS

```
RESULT 1
US-10-027-632-119945/c
; Sequence 119945, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 119945
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-119945
```

Query Match 3.7%, Score 82.8; DB 7; Length 1007;  
Best Local Similarity 58.0%; Pred. No. 3.1e-12;  
Matches 189; Conservative 0; Mismatches 127; Indels 10; Gaps 2;

```
OY 896 cccaggtggcgaactcttcctgaaagacccgacgagacatttaggtctgtg 955
    || || || || || || || || || || || || || || || || || || ||
DB 655 CCAGGGCTGAGTTAACTTTTCTGTAAGACCCAGAGTAATATTTTGGCTTTGAG 596
OY 956 gccatattgtctcaagtcacactactactctctgctctgtagcagaaagcaatagca 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 595 gccatattgtctctctgctgacatctattcaacttctgtttgacgcaaatgacccatag 536
OY 1016 acaatatgtcacaacaacatattggaaccatgtaaacatttattatattgatacggaaa 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 535 ATTAATGATTAATGAATGACATGCTGTTCCAAATAAACT--TTATTATGACACTGAAA 478
```

QY 1076 ccgaataatgtagctctcttttgatcttttcccaatcatataaagaatgaataaac 1135

Db 477 TTTTAAATTTTATATATATTTCTTTTTTGGATTTTGAACCATTTAAAAATTAAGAAACCA 418

QY 1136 tctatggtcgcaaggtctaaagcatctctcagctcagtcgagcgcggaatttgagctgt 1195

Db 417 TTTCTTTTCTTACACGGCTATGCCA-----TACAGTCACCAAGCCAGATTTTGCTCAT 366

QY 1196 gacctacagtcggcacaatccctgatt 1221

Db 365 AGGCGACAGTATGCCAACCCCTGTTT 340

```

RESULT 2
US-10-027-632-119946/c
Sequence 119946, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Many, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027.632
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIORITY APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-23
PRIORITY APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIORITY APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 119946
LENGTH: 1007
TYPE: DNA
ORGANISM: Human
US-10-027-632-119946

```

Query Match	3.7%	Score 82.8	DB 7	Length 1007
Best Local Similarity	58.0%	Pred. No. 3,1e-12		
Matches 189	Conservative	0	Mismatches 127	Indels 10
			Gaps	2
QY	896	cccaagttggaacaactcttcctgtaaagaacagacagaaactattttagctctglt	955	
Db	655	CCAGGGGCTCAGTTAACTTTTCTGTAAGAAGCCACAGAGTAAATATTTCGGCTTTGTAG	596	
QY	956	ggcactatggtctacagtcacaacactacatctctgcctctgtgagcaggaagcaatagca	1015	
Db	595	GCCATTAAAGTGCTCTGTGGCATCTATTCAACTTTCGTGTGTGCAGGCAAAATGCCACCTAG	536	
QY	1016	acacatctgcaacaacaacatactggaaccccatgaaacttlatatctatgatacggaa	1075	
Db	535	ATAATACATTAATGAATGACATGTGTCCAAATAAACT--TTTATTATGACACCTGAAA	478	
QY	1076	ccgaaataatagtctcttcttcttgatcttcttcccaacatataaanaacglaaanaactac	1135	
Db	477	TTTTAAATTCATATTAATCTTTTTTGGATATTTTGAACCAATTAATAAAGAACCA	418	
QY	1136	tctatagtcgcaaggtctaaagcaatcttcagcttagcagtcgagcgtgatttggctctgt	1195	
Db	417	TTCTTTTGCTCAGAGGCTATGGCA-----TTAACGTCAACAGCCAGATTTTGGCTCAT	366	
QY	1196	gaactacagtttggcacaatcccttgatt	1221	

```

DB          365  AGGCCACAGTATGCGCAACCCCTGTT 340

RESULT      3
US-10-105-299-12107/c
; Sequence 12107: Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: P5950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Paim
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12107
; LENGTH: 37437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12107

```

Query Match	3.7%	Score	81.8	DB 7	Length	37437
Best Local Similarity	62.9%	Pred. NO.	3.1e-11			
Matches	144	Conservative	0	Mismatches	82	Indels
						Gaps
						1:
QY 807	actctccctctctctctctatccatctgctgctggagcagttatctctctccaccatctca	866				
DB 32330	AGTCACGTTCTACTTCTTTCACAGTCTACTTCTTCCAGTATATATCTCTATTAACACATTA	32271				
QY 867	ca-----ctctctacaccgcttcccatctcgcgcgcagcttgagaaactcttccttaaa	923				
DB 32270	AAATGTCCTTTTGCAATAGTCACAGATATCTTACACAGGGCTCAGCAAACTTTTCTGAAA	32211				
QY 924	gaaccagacaggaactattttagtgcctctgtctgtgcacatcttgctctcaatcacaactactca	983				
DB 32210	GAGCCAGATGCTAATATATTTTATAGCTTCTTGCGCCAGCATCTCTCTGTGCACATATCA	32151				
QY 984	ctctcgcctctgttagcaggaagaacttagcacacaatalctgtcaaaac	1032				
DB 32150	TTTTCTGCTGCTATATAGTACAAACACACTTATTAACAAATATAGTAAACACATC	32102				

```

RESULT      4
US-10-105-299-12108/c
; Sequence 12108, Application US/10105299
; GENERAL INFORMATION:
APPLICANT: Rosen, et. al
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS950
CURRENT APPLICATION NUMBER: US/10/105,299
CURRENT FILING DATE: 2002-03-26
NUMBER OF SEQ ID NOS: 15197
Prior Application removed - See File Wrapper or Paim
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12108
LENGTH: 37442
TYPE: DNA
ORGANISM: Homo sapiens
US-10-105-299-12108

Query Match          3 6%; Score 81.4; DB 7; Length 37442:
Best Local Similarity 63.0%; Pred. No. 3.9e-11;
Matches 145; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

QY    809   tctctcttcttcattcaccacagctgtaggcagttaacctgtcccccacgctaca 868
       |||  ||||| ||||| |||  |||  |||  |||  |||  |||  |||  |||
Db    32329 TCACATTCATTCTTTCATGTGTCACACTCTTTCCAGATATATTTCTTAATAACAACATMAAA 32270

QY    869   ---ttctactccgglttccccatgctcggaaccagggttgggcaaacctccctgtaaga 925
       |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    32269 ATGTGCTCCTTTGCANATGGCTCAGATATCTTAACACAGGGGTCAACCAACTTTTCTCGAANAAGA 32210

```







```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7748
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-7748
```

```
Query Match 3.5%; Score 78.2; DB 7; Length 754;
Best Local Similarity 64.1%; Pred. No. 5.1e-11;
Matches 116; Conservative 1; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 892 tggaccaggttggcaaacctctctgttaagaacagaggaactatttagctct 951
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
DB 239 TGGAGCAGGCGCTTGGCAAACTTTCTGTAAAGGCAAGATGGTAATATTATAGGCTTT 180
QY 952 gtgtccatattgtctcaagtcacactactcatctctgtctgttagcagaagaatt 1011
    ||| | | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 179 GTGACCTTTCAGCGCTCTGTGACGACTACTCAACTCTGCGATTGTGTGAAGAAACAGCC 120
QY 1012 agcaacaatattgtcaacaacatatgtgaccccaatgaacattattattatgtagatg 1071
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
DB 119 ACAGACAAATACGAAAAAAGAAACATGTTGATTCACAACTTTATTATTATAAACA 60
QY 1072 g 1072
DB 59 G 59
```

```
RESULT 11
US-10-104-047-364
; Sequence 364, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NOVEL FULL LENGTH CDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 364
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-364
```

```
Query Match 3.4%; Score 77.2; DB 7; Length 2960;
Best Local Similarity 74.6%; Pred. No. 1.8e-10;
Matches 97; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

DB 2470 ggaagcagggctcagcaaacatttctgttaagggccagataagataatttgggcttgg 2529
QY 953 ttgtccatattgtctcaagtcacactactcatctgtcctctgttagcagaagaacatta 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2530 tgggacctatgtctctgtcaacaagtcactctgtctgttaacaaaaagcagctca 2589
QY 1013 gcaacaatat 1022
```

```
DB 2590 gacaatatgt 2599
```

```
RESULT 12
US-10-027-632-186700/c
; Sequence 186700, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186700
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-186700
```

```
Query Match 3.4%; Score 76.8; DB 7; Length 575;
Best Local Similarity 63.3%; Pred. No. 1.1e-10;
Matches 133; Conservative 1; Mismatches 73; Indels 3; Gaps 1;
```

```
QY 891 ctggaccaggttggcaaacctctctgttaagaacagaggaactatttagctc 950
    ||||| || ||||| ||||| ||||| ||||| || ||||| || ||||| ||
DB 384 CTGAGCAGGAGAGTAAAGCACTTTCTGTAAAGGCGAGATGATATTTTAGGATT 325
QY 951 ttgtccatattgtctcaagtcacactactcatctctgtcctctgttagcagaagaacatt 1010
    || || || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 324 TGGAGCTTTCAGCGCTCTGTGACGACTACTCAACTCTGCGATTGTGTGAAGAAATCAGC 265
QY 1011 tagcaacaatattgtcaacaacatatgtgaccccaatgaacattattattatgtagatg 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 264 CACAGATATATATATAAACAATGAGCATGCGTGT---TCCAGTAACACTTATGTTCAC 208
QY 1071 ggaaccgtgaaaaatatgtctctcttttga 1100
DB 207 TGAATCTGATTTTACACGTTTCATGTCA 178
```

```
RESULT 13
US-10-027-632-142654/c
; Sequence 142654, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142654
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142654
```

```
Query Match          3.4%; Score 76.6; DB 7; Length 538;
Best Local Similarity 64.7%; Pred. No. 1.2e-10;
Matches 130; Conservative 0; Mismatches 69; Indels 2; Gaps 1;
```

```
QY 901 gttggcaaacctctctctgtaagaacacagacaggaacatttaggcctgtgtgcat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 214 GGTGGCAAACTTTCTCTTAAAGGACGAGTAGTAACATTTTAGGTTTGCAGGCAC 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atgctcagtcacacactcactcctcctcgtctgtagcagaaagcaattagcaaat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 154 ATGCTCTGTAGCACTACTACTCCTCTGCACTAGCCTCAAGCACCACAGACAT 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacatagtaccccaatgaacttattt--attatgatacggaaacct 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 94 ATTTAATTAGCGCAACGTCGCGCTCTCCATTAAACTTCATTGTTGTACTGAATTT 35
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1079 gaaatacatgtctctctctg 1099
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 34 GAAATTTATATATTTCATG 14
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 14
; US-09-785-276A-55307/c
; Sequence 55307, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55307
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-785-276A-55307
```

```
Query Match          3.4%; Score 76.6; DB 5; Length 608;
Best Local Similarity 74.0%; Pred. No. 1.3e-10;
Matches 97; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 901 gttggcaaacctctctctgtaagaacacagacaggaacatttaggcctgtgtgcat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 GTTGGCAAACTTTCTCTTAAAGGACGAGTAGTAACATTTTAGGTTTGCAGGCAT 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atgctcagtcacacactcactcctcctcgtctgtagcagaaagcaattagcaaat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 AAGGCTCTTGTCCAACTATTCATCTCTGCACTTTTGTGAGAAAGCAGCAGTACAT 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacaa 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 GTATAACAAA 53
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 15
; US-10-027-632-233742/c
; Sequence 233742, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233742
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-233742
```

```
Query Match          3.4%; Score 75.6; DB 7; Length 556;
Best Local Similarity 65.9%; Pred. No. 2.3e-10;
Matches 108; Conservative 1; Mismatches 55; Indels 0; Gaps 0;
```

```
QY 901 gttggcaaacctctctctgtaagaacacagacaggaacatttaggcctgtgtgcat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 472 GTTGGCAAACTTTCTCTTAAAGGACGAGTAGTAACATTTTAGGTTTGCAGGCAT 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atgctcagtcacacacactcactcctcctcgtctgtagcagaaagcaattagcaaat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 412 ATGCTCTGTATCACAACACTGTGCTCAGCCACTGTAGCAAAAAGCATTTGTAAACAT 353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacatagtaccccaatgaacttatttattat 1064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 352 ATGTAAGTGAATGAACATGGGTGCTTCCAAATAAATTTAATTT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Search completed: August 16, 2002, 05:57:46
Job time: 11760 sec
```



1

2

3

4

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 07:54:01 ; Search time 5721.79 Seconds  
(without alignments)  
8469.757 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatcgcctcactatagg.....aaatcccaagaagcttgc 2240

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1:	/cgn2_6/ptodata/1/pna/US086.COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/US06.COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US08.COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US081.COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US082.COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US083.COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US084.COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US085.COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US086.COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US087.COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US088.COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US089.COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US090.COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US091.COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US092.COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US093.COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US094.COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US095.COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US096.COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US097.COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US098.COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US099.COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US100.COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US101.COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US102.COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US103.COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US104.COMB.seq:*
29:	/cgn2_6/ptodata/1/pna/US105.COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US106.COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US107.COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US108.COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US109.COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US110.COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US111.COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US112.COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US113.COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US114.COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US115.COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US116.COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US117.COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US118.COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US119.COMB.seq:*

Result No.	Score	Query length	ID	Description
1	2240	100.0	2286 19 US-09-515-369B-1	Sequence 1, Appl
2	420	18.8	46 19 US-09-528-409-108247	Sequence 108247,
3	420	18.8	35 US-09-933-524-108247	Sequence 108247,
4	420	18.8	46 35 US-09-933-524-108247	Sequence 108247,
5	93	4.2	338 1 PCT-US01-15674A-421	Sequence 421, App
6	93	4.2	338 22 US-09-573-080A-421	Sequence 421, App
7	87	3.9	196093 63 US-60-245-228-200	Sequence 200, App
8	87	3.9	196093 64 US-60-258-227-80	Sequence 80, Appl
9	86	3.8	224 1 PCT-US01-15674A-420	Sequence 420, App
10	86	3.8	224 22 US-09-573-080A-420	Sequence 420, App
11	86	3.8	76534 63 US-60-243-468-358	Sequence 358, App
12	83.2	3.7	179040 61 US-60-226-176-1401	Sequence 1401, Ap
13	83.2	3.7	179040 62 US-60-233-468-1401	Sequence 1401, Ap
14	83.2	3.7	179040 70 US-60-313-371-1401	Sequence 1401, Ap
15	83	3.7	153174 60 US-60-212-664-256	Sequence 256, App
16	82.8	3.7	1007 24 US-09-634-306B-119945	Sequence 119945,
17	82.8	3.7	1007 60 US-60-212-358-69	Sequence 69, Appl
18	82.4	3.7	17947 60 US-09-070-694-1633	Sequence 1633, Ap
19	81.8	3.7	288 14 US-60-044-030-1633	Sequence 1633, Ap
20	81.8	3.7	288 43 US-60-172-360-1441	Sequence 1441, Ap
21	81.8	3.7	511 56 US-60-185-216-346	Sequence 346, App
22	81.8	3.7	37437 1 PCT-US01-01354-38977	Sequence 38977, A
23	81.8	3.7	37437 30 US-09-764-905-38977	Sequence 9956, Ap
24	81.8	3.7	37437 35 US-09-950-083-9956	Sequence 9956, Ap
25	81.8	3.7	37437 37 US-10-092-399-38977	Sequence 38977, A
26	81.8	3.7	37437 37 PCT-US01-01354-38977	Sequence 38977, A
27	81.4	3.6	37442 30 US-09-764-905-38978	Sequence 38978, A
28	81.4	3.6	37442 35 US-09-950-083-9957	Sequence 9957, Ap
29	81.4	3.6	37442 37 US-10-092-399-38978	Sequence 38978, A
30	81.4	3.6	37442 37 US-10-092-399-38978	Sequence 38978, A
31	81.2	3.6	161395 61 US-60-226-176-1603	Sequence 1603, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

```

C 32 81.2 3.6 161395 62 US-60-233-468-1603 Sequence 1603, Ap
C 33 81.2 3.6 161395 70 US-60-313-371-1603 Sequence 1603, Ap
C 34 80.8 3.6 32768 59 US-60-208-965-68 Sequence 68, Appl
C 35 79.4 3.5 503 16 US-09-235-076-12893 Sequence 12893, A
C 36 79.4 3.5 503 16 US-09-248-797-34353 Sequence 34353, A
C 37 79.4 3.5 503 17 US-09-332-782-12893 Sequence 12893, A
C 38 79.4 3.5 503 29 US-09-737-223-12893 Sequence 12893, A
C 39 79.4 3.5 503 34 US-09-925-564-34353 Sequence 34353, A
C 40 79 79 749 24 US-09-634-306B-151080 Sequence 151080,
C 41 79 79 749 24 US-09-634-306B-151081 Sequence 151081,
C 42 78.4 3.5 312283 1 PCT-US00-27620-91 Sequence 91, Appl
C 43 78.2 3.5 754 24 US-09-634-306B-7748 Sequence 7748, Ap
C 44 78.2 3.5 118794 63 US-60-243-468-310 Sequence 310, Appl
C 45 78.2 3.5 118794 63 US-60-243-742-53 Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-515-369B-1
; Sequence 1, Application US/09515369B
; GENERAL INFORMATION:
; APPLICANT: Paul B. Fisher
; APPLICANT: Malavi T. Medireddi
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; FILE REFERENCE: 34611 070050.1685
; CURRENT APPLICATION NUMBER: US/09/515, 369B
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Human
US-09-515-369B-1

Query Match 100.0%; Score 2240; DB 19; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C 32 81.2 3.6 161395 62 US-60-233-468-1603 Sequence 1603, Ap
C 33 81.2 3.6 161395 70 US-60-313-371-1603 Sequence 1603, Ap
C 34 80.8 3.6 32768 59 US-60-208-965-68 Sequence 68, Appl
C 35 79.4 3.5 503 16 US-09-235-076-12893 Sequence 12893, A
C 36 79.4 3.5 503 16 US-09-248-797-34353 Sequence 34353, A
C 37 79.4 3.5 503 17 US-09-332-782-12893 Sequence 12893, A
C 38 79.4 3.5 503 29 US-09-737-223-12893 Sequence 12893, A
C 39 79.4 3.5 503 34 US-09-925-564-34353 Sequence 34353, A
C 40 79 79 749 24 US-09-634-306B-151080 Sequence 151080,
C 41 79 79 749 24 US-09-634-306B-151081 Sequence 151081,
C 42 78.4 3.5 312283 1 PCT-US00-27620-91 Sequence 91, Appl
C 43 78.2 3.5 754 24 US-09-634-306B-7748 Sequence 7748, Ap
C 44 78.2 3.5 118794 63 US-60-243-468-310 Sequence 310, Appl
C 45 78.2 3.5 118794 63 US-60-243-742-53 Sequence 53, Appl
```

```

Db 421 gctgtctctgtccctccacatccattatccagcactctcactgtgtcttcttctg 480
Oy 481 tctcatcactgtccctcttgacatcttattatccatagtagttagttaggttcttgtaat 540
Db 481 tctcatcactgtccctcttgacatcttattatccatagtagttagttaggttcttgtaat 540
Oy 541 gccccaatcccaatcagtggtggaaggggaggtgtggtggaaggaagtgccgtctggggcgtc 600
Db 541 gccccaatcccaatcagtggtggaaggggaggtgtggtggaaggaagtgccgtctggggcgtc 600
Oy 601 gctcactctctgtaggttagagactcgcgcctcccaagaacagatcttagctgtgtgca 660
Db 601 gctcactctctgtaggttagagactcgcgcctcccaagaacagatcttagctgtgtgca 660
Oy 661 gctatagccaagcagactgtgtgccaaggtatgccaagagatattgtttgtcttaagaa 720
Db 661 gctatagccaagcagactgtgtgccaaggtatgccaagagatattgtttgtcttaagaa 720
Oy 721 aataaacaacactgtgtagttagagtaggaggggtgtgtgtgtccagagagattggaa 780
Db 721 aataaacaacactgtgtagttagagtaggaggggtgtgtgtgtccagagagattggaa 780
Oy 781 gagtctgccaaagggtgtgttctactcaactctcctcttcttcttcaatcccaactagctgtg 840
Db 781 gagtctgccaaagggtgtgttctactcaactctcctcttcttcttcaatcccaactagctgtg 840
Oy 841 aggcagttatctgtgtcccccagcgtacactctcactcccggttcccaatgtcctggaccag 900
Db 841 aggcagttatctgtgtcccccagcgtacactctcactcccggttcccaatgtcctggaccag 900
Oy 901 gttggccaacactcttcctgttaagaagacaagaaactatttaggtctgtgtgcat 960
Db 901 gttggccaacactcttcctgttaagaagacaagaaactatttaggtctgtgtgcat 960
Oy 961 atgtgtcagtcacaactatcactatctctgtcctgtgtgacgacgaagaaatagacaact 1020
Db 961 atgtgtcagtcacaactatcactatctctgtcctgtgtgacgacgaagaaatagacaact 1020
Oy 1021 atgtcaacaacatagttagaccatgaacaaactatttattatgtagtagaagaaactgtg 1080
Db 1021 atgtcaacaacatagttagaccatgaacaaactatttattatgtagtagaagaaactgtg 1080
Oy 1081 aaataatgtcttcttcttgatlttttccccaatcatataaanaagcgtaaanaactactctta 1140
Db 1081 aaataatgtcttcttcttgatlttttccccaatcatataaanaagcgtaaanaactactctta 1140
Oy 1141 ggtgcgaaggttaagcactcttcagcttagcaggtgtgacgactgtgattgtgtgtgact 1200
Db 1141 ggtgcgaaggttaagcactcttcagcttagcaggtgtgacgactgtgattgtgtgtgact 1200
Oy 1201 acagttggccaatccctgtatccccaanaatgtatccctcagggagtgtggccaataacttat 1260
Db 1201 acagttggccaatccctgtatccccaanaatgtatccctcagggagtgtggccaataacttat 1260
Oy 1261 gggaggtgtgtgattaaacaagagtttaagaaagcatcagacatltccaggaagcgttagcac 1320
Db 1261 gggaggtgtgtgattaaacaagagtttaagaaagcatcagacatltccaggaagcgttagcac 1320
Oy 1321 atggcagggcctcttaactgaacctgaacctatgttattatgttagtagaatactttcagag 1380
Db 1321 atggcagggcctcttaactgaacctgaacctatgttattatgttagtagaatactttcagag 1380
Oy 1381 acaagaattctcctcaaaacttagagtcctgtagagactgtgtgtggaacaacagcctctgtg 1440
Db 1381 acaagaattctcctcaaaacttagagtcctgtagagactgtgtgtggaacaacagcctctgtg 1440
Oy 1441 atgcccctcactgtggcacatgtgtagaaatctagagctgtgtgtgtgtgtgtgtgtgtgtgt 1500
Db 1441 atgcccctcactgtggcacatgtgtagaaatctagagctgtgtgtgtgtgtgtgtgtgtgtgt 1500
Oy 1501 gtccagagctctttagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
Db 1501 gtccagagctctttagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
```

```
QY 1561 ggttcagagccacacagcccaagactagcccccacagacccagcctccagcttctc 1620
Db 1561 ggttcagagccacacagcccaagactagcccccacagacccagcctccagcttctc 1620
QY 1621 ggttcagagccacacagcccaagactagcccccacagacccagcctccagcttctc 1680
Db 1621 ggttcagagccacacagcccaagactagcccccacagacccagcctccagcttctc 1680
QY 1681 atttggatgagatataaacccttaggacatggtagctcagtcaggaacaacac 1740
Db 1681 atttggatgagatataaacccttaggacatggtagctcagtcaggaacaacac 1740
QY 1741 ctcaactctccgcccggatgaatgaagagagagagagagagagagagagagagagag 1800
Db 1741 ctcaactctccgcccggatgaatgaagagagagagagagagagagagagagagagag 1800
QY 1801 ggaactctgcccagcagatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1860
Db 1801 ggaactctgcccagcagatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1860
QY 1861 tgggtatgctgcacagtaataatgataatgataatgataatgataatgataatgata 1920
Db 1861 tgggtatgctgcacagtaataatgataatgataatgataatgataatgataatgata 1920
QY 1921 ggtgtcttctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1980
Db 1921 ggtgtcttctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1980
QY 1981 agactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
Db 1981 agactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
QY 2041 tgaactcaaaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
Db 2041 tgaactcaaaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
QY 2101 cttaaatgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtag 2160
Db 2101 cttaaatgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtag 2160
QY 2161 cttcttcacactcctcttcttcttcttcttcttcttcttcttcttcttcttct 2220
Db 2161 cttcttcacactcctcttcttcttcttcttcttcttcttcttcttcttcttct 2220
QY 2221 aaatccacaaagccttgc 2240
Db 2221 aaatccacaaagccttgc 2240

RESULT 2
US-09-528-409-108247/c
: Sequence 108247, Application US/09528409
: GENERAL INFORMATION:
: APPLICANT: Drmanac, Radoje T.
: APPLICANT: Labat, Ivan
: APPLICANT: Stache-Crain, Birgit
: APPLICANT: Dickson, Mark
: APPLICANT: Jones, Lee W.
: TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
: TITLE OF INVENTION: From Various Libraries
: FILE REFERENCE: 774
: CURRENT APPLICATION NUMBER: US/09/528,409
: CURRENT FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/125,453
: PRIOR FILING DATE: 1999-03-19
: NUMBER OF SEQ ID NOS: 116231
: SOFTWARE: Hy-patent.pl Version 3.1
: SEQ ID NO 108247
: LENGTH: 426
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-528-409-108247
```

```
Query Match 18.8%; Score 420; DB 19; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ggttggaaggggagtgaggagagagagagagagagagagagagagagagagagag 615
Db 425 ggttggaaggggagtgaggagagagagagagagagagagagagagagagagagag 366
QY 616 gtaagactcggccctccaggaacaaagatcagagctggtagcagctatagccaagag 675
Db 365 gtaagactcggccctccaggaacaaagatcagagctggtagcagctatagccaagag 306
QY 676 actgctggcagagagatgtaaaagagatattgttcttaagaaataaacaacatga 735
Db 305 actgctggcagagagatgtaaaagagatattgttcttaagaaataaacaacatga 246
QY 736 gtaagatgagagagagagagagagagagagagagagagagagagagagagagag 795
Db 245 gtaagatgagagagagagagagagagagagagagagagagagagagagagagag 186
QY 796 ggtttactactcctcctcttcttcttcttcttcttcttcttcttcttcttcttct 855
Db 185 ggtttactactcctcctccttcttcttcttcttcttcttcttcttcttcttcttctt 126
QY 856 ccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 915
Db 125 ccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 66
QY 916 cctgtaagaaacacagagagagagagagagagagagagagagagagagagagag 975
Db 65 cctgtaagaaacacagagagagagagagagagagagagagagagagagagagag 6

RESULT 3
US-09-933-524-108247/c
: Sequence 108247, Application US/09933524
: GENERAL INFORMATION:
: APPLICANT: Drmanac, Radoje T.
: APPLICANT: Labat, Ivan
: APPLICANT: Stache-Crain, Birgit
: APPLICANT: Dickson, Mark
: APPLICANT: Jones, Lee W.
: TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
: TITLE OF INVENTION: From Various Libraries
: FILE REFERENCE: 774
: CURRENT APPLICATION NUMBER: US/09/933,524
: CURRENT FILING DATE: 2001-08-20
: PRIOR APPLICATION NUMBER: 09/528,409
: PRIOR FILING DATE: 2000-03-17
: NUMBER OF SEQ ID NOS: 116231
: SOFTWARE: Hy-patent.pl Version 3.1
: SEQ ID NO 108247
: LENGTH: 426
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-933-524-108247

Query Match 18.8%; Score 420; DB 35; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ggttggaaggggagtgaggagagagagagagagagagagagagagagagagagag 615
Db 425 ggttggaaggggagtgaggagagagagagagagagagagagagagagagagagag 366
QY 616 gtaagactcggccctccaggaacaaagatcagagctggtagcagctatagccaagag 675
Db 365 gtaagactcggccctccaggaacaaagatcagagctggtagcagctatagccaagag 306
QY 676 actgctggcagagagatgtaaaagagatattgttcttaagaaataaacaacatga 735
```

```
Db 305 ACTGCTGGCCAGGAGTTCGAAAGACTATTTTGTCTTAAGAAATTAACACACTCA 246
    |||
Qy 736 gttctgagatgagagagagggggtgtctgctgcagagagattgggaagatctgcgaaggt 795
    |||
Db 245 GATATGAGATGAGAGGGGGGTGTGTGTCAGAGAGATTTGGAAAGAGTCTGCCAAGGGGT 186
    |||
Qy 796 gtgtctactcaactctctctcttcttctatctccactcaagctggagagattatctct 855
    |||
Db 185 GTTTTACTCACTCTCTCTTTCTTTTATCTCCACTAGCTGGAGGAGATATCTGT 126
    |||
Qy 856 ccccaagctcaactctctctctcccgcttccatgctgagcccaagcttgggcaactct 915
    |||
Db 125 CCCCACGCTCACATTCCTACTCCCGTTTCCATGCTGAGCCAGGAGTGGGCAAACTCTT 66
    |||
Qy 916 ccgttaagaagacagagagaaactattttagctctgtgtgcatatgtgtccagtcaca 975
    |||
Db 65 CCTGTAAAGAACACAGAGAACTATTTTAGCTCTGTGTCCATATGTGTCTACATGACA 6
```

```
RESULT 4
US-09-933-524A-108247/c
; Sequence 108247, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Dimaac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-Patent.pl Version 3.1
; SEQ ID NO 108247
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524A-108247
```

```
Query Match 18.8%; Score 420; DB 35; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 556 ggttggaagggggggggtggggaagagtgctgctggtggtctgtgctactctcgagg 615
    |||
Db 425 GGTGGGAAGGGGGGAGTGGGAGAGAGTGGCGCTGTGGGCTGTCTACTTCTGGAGG 366
    |||
Qy 616 gtaagaactcggccctccagagacaaagatcaagctcgtgagagctatagccaagcag 675
    |||
Db 365 GTAAGACTCGGGCCCTCCAGGAACAAAGATTTCAGAGCTGGTGCGAGCTATAGCCAGCAG 306
    |||
Qy 676 actgctggccaaggaatgcgaagagatattgttctgcttaagaataaacaactcga 735
    |||
Db 305 ACTGCTGCCACGAGTATGCAAAAGAGTATTTTGTCTTAAGAAATTAACACACTCA 246
    |||
Qy 736 gtaagaatgagagaggggggtgtgtgctgcagagagattgggaagatctgcgaaggt 795
    |||
Db 245 GATATGATGAGAGGGGGGTGTGTGTCAGAGAGATTTGGGAAGAGTCTGCCAAGGGT 186
    |||
Qy 796 gtgtctactcaactctctcttcttctatctccactcaagctggagagattatctctgt 855
    |||
Db 185 GTTTTACTCACTCTCTCTTTCTTTTATCTCCACTAGCTGGAGGAGATATCTGT 126
    |||
Qy 856 ccccaagctcaactctctctccgcttccatgctgagcccaagcttgggcaactct 915
    |||
Db 125 CCCCACGCTCACATTCCTACTCCCGTTTCCATGCTGAGCCAGGAGTGGGCAAACTCTT 66
    |||
```

```
Qy 916 cctgtaagaagacagagaaactattttagctctgtgtccatatgtgtcagtcaca 975
    |||
Db 65 CCTGTAAAGAACACAGAGAACTATTTTAGCTCTGTGTCCATATGTGTCTACATGACA 6
    |||
RESULT 5
PCT-US01-15674A-421/c
; Sequence 421, Application PC/TUS0115674A
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: PCT/US01/15674A
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 421
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(338)
; OTHER INFORMATION: cheshire_b
PCT-US01-15674A-421
```

```
Query Match 4.2%; Score 93; DB 1; Length 338;
Best Local Similarity 65.8%; Pred. No. 4.4e-15;
Matches 177; Conservative 1; Mismatches 66; Indels 25; Gaps 2;
```

```
Qy 900 gttgggaacactctctctgtaagaacacagagagaaactattttagctctgtgtcca 959
    |||
Db 335 GGGTCGGAACCTTTTCTGTAAAGGCGCAGATGTAATTTTACGCTTGGCGGCCCA 276
    |||
Qy 960 tatgtctcagtcacaactactcatctctgctctgtagcacgaagaacttagcaaca 1019
    |||
Db 275 TACGCTCTCTGCGCACTACTCAACTCTGCGTGTGAGGCAACACAGCATAGACAA 216
    |||
Qy 1020 tatgtcaacaacataatgtgaccccat--gaaacttattattatgataaggaacc 1077
    |||
Db 215 TATATAAACGAATGGCGCTGCTGTGTCATTAACCTTATTTATGAGACACTGAATTT 156
    |||
Qy 1078 tga-----aaatagtcttctctttagatttttcccaatc 1114
    |||
Db 155 TGAATTTCAATTAATTTTCACATGTCACAAATATVCTCTTTTGATTTTCTTCCAAAC 96
    |||
Qy 1115 attaaaaacgtlaaaactactcttagt 1143
    |||
Db 95 ATTTAAAAATGTAAAAACCAATTCCTAGTT 67
    |||
```

```
RESULT 6
US-09-573-080A-421/c
; Sequence 421, Application US/09573080A
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENER
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 421
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(338)
; OTHER INFORMATION: cheshire_b
```



```

; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichewicz, J; Mijosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; US-09-573-080a-421

```

```

Query Match 4.2%; Score 93; DB 22; Length 338;
Best Local Similarity 65.8%; Pred. No. 4.4e-15;
Matches 177; Conservative 1; Mismatches 66; Indels 25; Gaps 2;

```

```

QY 900 ggttgagcaactcttctctgaagaacacagagaactatttagctctgtgacca 959
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335 GGGTCGCAAACTTTCTGTAAGGCCAGATGTAATATTATTTAGCTTTCGGGCCA 276
QY 960 tatgtctcagtcacactctctctctctctctctctctctctctctctctct 1019
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 275 TAGGTCCTGTCGCAACTCTGCACTCTGCGTGTAGCGCAAAAGCAGCATAGCAA 216
QY 1020 tatgcacaacaacatagtaccccat--gaaaccttattatgatacgaac 1077
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 215 TATATAACGAATGCGGCTGCTGTGCCAATTAACCTTATTTATGACATGAAAT 156
QY 1078 tga-----aataatgctctctctctctctctctctctctctctctcc 1114
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 155 TGAATTCATATATATTTTTCACATGTCACAAATATATCTCTTTGATTTTCTTCAACC 96
QY 1115 attaaaaacgtaaaactactcttgat 1143
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 95 ATTTAAAAATGTAAAAACCATCTTAACT 67

```

```

RESULT 7
US-60-245-228-200
; Sequence 200, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000878
; CURRENT APPLICATION NUMBER: US/60/245,228
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200
; LENGTH: 196093
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(196093)
; OTHER INFORMATION: n = A,T,C or G
US-60-245-228-200

```

```

Query Match 3.9%; Score 87; DB 63; Length 196093;
Best Local Similarity 70.1%; Pred. No. 9e-12;
Matches 117; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 881 ttctcagcttgagccacagtggtggaacactctctctgaagaacagacagaa 940
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170687 ttgttaacctacatcagtgatggaacactctctctgaagaagcagacga 170746
QY 941 tttaagctctgtgtccataatggtctcagtcacaactactctctctctgtaga 1000

```

```

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170747 tttaagctcttgagccataatggtctcagtcacaactactctctgtaga 170806
QY 1001 cgaagaacttgacacatattgtaacaacacatattgaccccaty 1047
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170807 tgaaagcagccatagacataataatgaaatgagtggtgagcaty 170853

```

```

RESULT 8
US-60-258-272-80
; Sequence 80, Application US/60258272
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL0001041
; CURRENT APPLICATION NUMBER: US/60/258,272
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 196093
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(196093)
; OTHER INFORMATION: n = A,T,C or G
US-60-258-272-80

```

```

Query Match 3.9%; Score 87; DB 64; Length 196093;
Best Local Similarity 70.1%; Pred. No. 9e-12;
Matches 117; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 881 ttcccatgcttgagccacagtggtggaacactctctctgaagaacagacaga 940
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170687 ttgttaaacctacatcagtgatggaacactctctctgaagaagcagacga 170746
QY 941 tttaagctctgtgtccataatggtctcagtcacaactactctctctgtaga 1000
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170747 tttaagctcttgagccataatggtctcagtcacaactactctctgtaga 170806
QY 1001 cgaagaacttgacacatattgtaacaacacatattgaccccaty 1047
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170807 tgaaagcagccatagacataataatgaaatgagtggtgagcaty 170853

```

```

RESULT 9
PCT-US01-15674A-420/c
; Sequence 420, Application PC/TUS0115674A
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: PCT/US01/15674A
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 420
; LENGTH: 224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(224)
; OTHER INFORMATION: cheshire_a
PCT-US01-15674A-420

```

```

Query Match 3.8%; Score 86; DB 1; Length 224;

```



Db	5023	cttggagaagcttaggaacaggggtcaacaacactttttgtgaagaaaccagaagttaa	50823
QY	939	tattttaggcctctgtgtgcataatggltcagtcacactactcatctgtgcctgtag	998
Db	5083	tatttcaggcttttgcagggccataatgctctcctgtcgaactgtccaaacctgcaggtag	51428
QY	999	cacgaagacatttgcacacataatgtcaacaacataatgtgacccatgaaactttat	10568
Db	5143	caccaagggcgcgtctagaacata-ctcaagcaagaagcctgtgtcctataaataattat	52018
QY	1059	tattatgatgcggaagccgtgaataatgtctctctcttcttgattttt	11068
Db	5202	tatgaccggaatttgaatttcatactatcctttaaagtttctctttttt	52498

RESULT 13

```

US-60-233-468-1401
? Sequence 1401, Application US/60233468
? GENERAL INFORMATION:
? APPLICANT: Ring, Huijun Z.
? APPLICANT: Malsen, Careth
? APPLICANT: Townley, David
? APPLICANT: Morris, Macdonald
? APPLICANT: Valdes, Ana
? TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes
? FILE REFERENCE: GX-0013-2 P
? CURRENT APPLICATION NUMBER: US/60/233,468
? CURRENT FILING DATE: 2000-09-18
? NUMBER OF SEQ ID NOS: 2488
? SOFTWARE: PERL Program
? SEQ ID NO 1401
? LENGTH: 179040
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc-feature
? OTHER INFORMATION: GB:AC022732_3
US-60-233-468-1401

```

Query Match	3.7%	Score 83.2	DB 62	Length 179040
Best Local Similarity	63.2%	Prod. NO. 1e-10		
Matches 144	Conservative 0	Mismatches 83	Indels 1	Gaps 1
QY 879	cgttcccatgcctcgcgaccaggtctgggcaaaccttccctgtaaaagacacagagaac	938		
Db 5023	ctttgagcaagccttagcgacaggggtcacaacaaccttttttgaagaaccacgaagaa	5082		
QY 939	tatttttaggcctcgtgctgcatactgtcttcgctcaacatactatctctgcctctgag	998		
Db 5083	tatttcaggtcttgcgaagcacaatggtctctgcgcgaactgctcaacctgcgcgaagtg	5142		
QY 999	cacgaaagaacattagcaacaatattgcacaacaacatattggaacccatgaaaccttatt	1058		
Db 5143	caccaagcgactcatagacaata-ctcaagcaagaagctgctgctataaataatttatt	5201		
QY 1059	tatttgatagatcggaacctgaataataatgctctctctttgattttt	1106		
Db 5202	tatgactgaaatttggaaatttcataataactttatgctttctctttttt	5249		

## RESULT 14

```

US-60-313-371-1401
: Sequence 1401, Application US/60313371
:
: GENERAL INFORMATION:
:
: APPLICANT: Ring, Huijun Z.
:
: APPLICANT: Malsen, Gareth
:
: APPLICANT: Townley, David
:
: APPLICANT: Morris, Macdonald
:
: TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated with ADME Genes
:
: FILE REFERENCE: GX-0013-5 P
:
: CURRENT APPLICATION NUMBER: US/60/313-371

```

```

? CURRENT FILING DATE: 2001-08-16
? NUMBER OF SEQ ID NOS: 2447
? SOFTWARE: PERL Program
? SEQ ID NO 1401
? LENGTH: 179040
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: GB:AC022732_3
US-60-313-371-1401

```

Query Match	3.7%	Score 83.2	DB 70	Length 179040
Best Local Similarity	63.2%	Pred. No. 1e-10		
Matches 144	Conservative 0	Mismatches 83	Indels 1	Gaps 1
QY 879	cgttcccatgccttgcagccaggcttggcgaac	ctcttcgttaagaaccagacaggaac	938	
DB 5023	cttttagcaagagcttagaagaagggtcaacaac	ctttttgttaagaaccagagaagtaaa	5082	
QY 939	tatttagagctctggtgcacatgctcagtcacaa	taactaactatctctgctctgtag	998	
DB 5083	tatttcaggcttttgcaggcacatggtctcgcgtc	gaactgtccaaacctgcgaaggtag	5142	
QY 999	cacgaagaacattgcacaanaatgacaaaca	atatygcacccatgaaaccttatt	1058	
DB 5143	caccaagagagcatagacaaca-c	ctcaagacaagagagtgctcaataaatttatt	5201	
QY 1059	tatttgatagcgaacactgnaataatgctttctt	tgattttttt	1106	
DB 5202	tatgcgcgaacttgcgaattctcatacttat	tgltttctttttt	5249	

## RESULT 15

```

US-60-212-664-256/c
; Sequence 256, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Spier, Gene
; APPLICANT: Greenberg, Simon
; APPLICANT: Rabin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO00687
; CURRENT APPLICATION NUMBER: US/60/212, 664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 256
; LENGTH: 163174
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ..(163174)
; OTHER INFORMATION: n = A,T,C or G
US-60-212-664-256

```

Query Match	3.7%	Score 83	DB 60	Length 163174
Best Local Similarity	65.2%	Pred. No. 1,1e-10		
Matches 122	Conservative 0	Mismatches 65	Indels 0	Gaps 0
QY	903	tgggaaacattctctgtaaaagaccacaacaggaacttlttagctctgtgtgcacatat	962	
Db	92341	TCGGCAACCTTCCCTGTAAAGGGTCACATACGTAAATATTTTAAAGGCTTGCGGCCACAG	92282	
QY	963	ggtctagtcacaactactcaactctctgacctgttagcaggaagcaattagcaacaatat	1022	
Db	92281	AGCTCTCTGTCCAAATGTGTCACACTTACTGTGTGGCCCAAAAGCAGGCATATAAATAAGCT	92222	

OY 1023 gccaacacacatactgtagccccaatgaacattatattatgatacggaaacctgaa 1082  
 |||||  
 Db 92221 GTPAATGAATTAACGTGSCAATGTTCACAAAAAATTATTTATGGAGAGTTGAATTCAT 92162  
 OY 1083 ataagc 1089  
 |||||  
 Db 92161 ATAAATT 92155

Search completed: August 16, 2002, 08:16:27  
 Job time: 16584 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 03:39:47 ; Search time 3499.63 seconds

(without alignments)  
8638.965 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatcagctcactatagg.....aatctccacaagccttgc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pin:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	81	3.6	480	10	BI262402 602953836
2	80.6	3.6	354	10	T58770 YB80H02.S1
3	80.4	3.6	390	10	I44398 HUMEST1E3.H
4	78.4	3.5	521	12	A0695956 HS_2160.A
5	78.2	3.5	435	9	AW183569
6	78.2	3.5	459	9	AW125442
7	78.2	3.5	628	12	A0378792 RPT11-15
8	77.4	3.5	393	3	AI128823
9	77.4	3.5	411	10	W76494
10	77.2	3.5	680	12	AG154552
11	77.2	3.4	400	9	AI926274
12	77	3.4	322	9	AI240516 qn12h02.x
13	77	3.4	386	9	AA443938 zv51c03.s
14	77	3.4	393	9	AA444117 zv51c03.r
15	77	3.4	417	9	AI208768
16	76.8	3.4	509	10	BI061153
17	76.6	3.4	327	12	B98848

18	76.2	3.4	499	12	A0277351	A0277351 CITB1-E1-
19	76.2	3.4	644	12	A0039307	A0039307 CIT-HSP-2
20	76.2	3.4	5529	12	AF101969	AF101969 AF101969
21	76	3.4	573	12	A0633901	A0633901 RPT-11-4
22	75.6	3.4	530	12	A0508112	A0508112 RPT-11-2
23	75.6	3.4	628	12	A0020407	A0020407 CIT-HSP-2
24	75.6	3.4	799	10	BG535645	BG535645 602563353
25	75.2	3.4	488	9	AV604853	AV604853 AV604853
26	74.6	3.3	348	12	A0035111	A0035111 CIT-HSP-2
27	73.6	3.3	640	12	AG093736	AG093736 Pan trogl
28	73.2	3.3	381	9	AM574982	AM574982 UT-HF-BNO
29	73.2	3.3	512	12	A0768179	A0768179 HS_3160.B
30	73	3.3	519	12	A0029411	A0029411 RPT11-39
31	72.8	3.2	302	9	A1559884	A1559884 th87a12.x
32	72.8	3.2	427	12	A0815705	A0815705 HS_5394.A
33	72.2	3.2	312	10	BG230513	BG230513 na14a06
34	72	3.2	416	9	AM835797	AM835797 QVA-LT001
35	72	3.2	764	10	BI915216	BI915216 603184745
36	71.8	3.2	2032	11	AF161549	AF161549 Homo sapi
37	71.4	3.2	794	10	BF983586	BF983586 602306575
38	71.2	3.2	327	9	A1344682	A1344682 qp09h04.x
39	71.2	3.2	363	9	A1358986	A1358986 q19h02.x
40	71.2	3.2	372	9	A1370636	A1370636 ta88g12.x
41	71.2	3.2	380	9	A1244892	A1244892 q19f05.x
42	71.2	3.2	380	9	A1298789	A1298789 qm91c02.x
43	71.2	3.2	388	9	AI291781	AI291781 qm75b05.x
44	71.2	3.2	393	9	AI214952	AI214952 qm31d12.x
45	71.2	3.2	394	9	A1439755	A1439755 t161a07.x

#### ALIGNMENTS

RESULT 1  
BI262402 480 bp mRNA linear EST 17-JUN-2001  
602953836T1 NIH\_MGC\_99 Homo sapiens CDNA clone IMAGE:5088033 3 ,  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 480)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA Library  
Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/MLMU at:  
<http://image.llnl.gov>  
Plate: LMC1846 row: g column: 10  
High quality sequence start: 27  
High quality sequence stop: 480.

#### FEATURES

source  
1..480  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5088033"  
/clone\_lib="NIH-MGC\_99"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph. Vector: pOTB7; site\_1: XhoI; site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GCCACGAC(G). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley), using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library.

Query Match	3.6%	Score 81:	DB 10:	Length 480;
Best Local Similarity	74.5%	Pred. No. 8.9e-10:		
Matches 102:	Conservative 0:	Mismatches 35:	Indels 0:	Gaps 0:
QY	894	gaccacaggctggygcaaacctctctcgtgtaaagacacagacgaactattctagctcgt	953	
Db	156	GACCAGCAATCATGAAAACCTTTCTTGTAAGGCCAGAGAGTAAGTATTTAGGTTTGT	215	
QY	954	gtgcacatttgtctcaagtcacaacactactcaatctctgctctgtagcagaaagcaattg	1013	
Db	216	GGGCAATATGGCTCTTGCGCAATTACTTAATCTTGCCATCATCTATACGAGAAAGCAGCTAC	275	
QY	1014	caacaaatgtcacaaca 1030		
Db	276	AGACAAATATGTATACCAA 292		

RESULT	2				
LOCUS	T58770				
DEFINITION	T58770	354 bp	mRNA	linear	EST 09-FEB-1995
ACCESSION	Y580H02.s1	Stratagene liver (#937224) Homo sapiens cDNA clone			
VERSION	T58770	IMAGE:77523 5', mRNA sequence.			
KEYWORDS	T58770.1	GI:660607			
SOURCE	EST.				
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 354)				
AUTHORS	Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,				
	Chisoso,S., Dietrich,N., Dubuque,T., Fawello,A., Gish,W., Hawkins				
	M., Hulman,M., Kucaba,T., Lucy,M., Le,N., Mards,E., Moore				
	B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,				
	Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Treviski,E.,				
	Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.				
TITLE	Generation and analysis of 280,000 human expressed sequence tags				
JOURNAL	Genome Res. 6 (9), 807-828 (1996)				
MEDLINE	97044478				
COMMENT	Contact: Wilson RK				

Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 2225  
 High quality sequence stops: 271 Source: IMAGE Consortium, LINTL This  
 clone is available royalty-free through LINTL ; contact the IMAGE  
 Consortium (info@image.lml.gov) for further information.  
 Insert Length: 2225 Std Error: 0.00  
 Seq Primer: -21ml3  
 High quality sequence stop: 271.  
 Location/Qualifiers  
 1..354  
 FEATURES  
 source

```

/organism="Homo sapiens"
/db_xref="GDB:497268"
/db_xref="taxon:9606"
/clone IMAGE:77523
/clone_id="Stratagene liver (#937224)"
/sex="male"
/dev_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: liver; vector: pBluescript SK; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo

```

dt. Hepatectomy from normal male caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCACGTTTTTTTTTTTTTTTT 3'."

Query Match	Similarity	Score	DB	Length
Best Local Similarity	74.3%	Pred. No. 1e-09;		
Matches	101;	Conservative	0;	Mismatches 35; Indels 0; Gaps 0;
Oy	894	gaccacagtggtggcaacatctctcgttaagaacagacagaaactatttagagctcgt	953	
Db	181	GACCGAGGATCAGAAAACTTTTCTGTAAAGGCGNAGACTAAGTATTATTAGGTTTGT	240	
Oy	954	gtgcacatgtgtcttaagtcacaactactcactctctgctctgtgacagaaagcaatga	1013	
Db	241	GGGGCATTTGGTCTCTTGTGGCAATTACTCTGCATCATGACGAAAGCAGCTAC	300	
Oy	1014	caacacatctgcaca	1029	
Db	301	AGACATATGTACGCA	316	

RESULT	3
L44398	
LOCUS	L44398 390 bp mRNA linear EST 17-JAN-1996
DEFINITION	HOMESTIE3 Human thymus NSTH II Homo sapiens cDNA, mRNA sequence.
ACCESSION	L44398
VERSION	L44398.1 GI:1048762
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 390)
AUTHORS	Lamerdin,J.E., Athwal,R.S., Patanjali,S., Weissman,S. and Carrano ,A.V.
TITLE	Chromosomal localization and expressed sequence tag generation of clones from a normalized human adult thymus cDNA library
JOURNAL	Genome Res. 5 (4), 359-367 (1995)
MEDLINE	96356441
COMMENT	Contact: Human Genome Center, Biology and Biotechnology Research Program Lawrence Livermore National Laboratory Livermore, CA 94550, USA.
FEATURES	Location/Qualifiers
source	1..390

clone_lib="human thymus NSTH II"	
/note="From adult human thymus NSTH II cells; randomly	
primed; normalized"	
BASE COUNT	118 a 85 c 64 g 123 t
ORIGIN	

	Query Match	3.68;	Score 80.4;	DB 10;	Length 390;
	Best Local Similarity	62.48;	Pred. No. 1.2e-09;		
	Matches 126;	Conservative	0;	Mismatches 76;	Indels 0;
				Gaps 0;	
QY	890	cttgagaccaggttgggcaacttctcttgaagaacccagacggaacttttagct	949		
DB	29	CCCTAGATCAAGAGTAGGAAACCTTTTCTGCAAGACAGATGTAACATTTTGCTCT	88		
QY	950	ctgtgtgccataatgtcttcagtcacaactactatctctgtcctcttgtagcagaagca	1009		
DB	89	TTGTAGGCAATATGTCCTCTCGCAATTACTCAACTTTTCCGCACTGTAGTGAGAAACCA	148		
QY	1010	tttagcaacaatattgcacaacaacatagtgagcccaatgaaaccttttattatgtata	1069		

Db 149 CCAAGACAATACATAAGCAAGGAGCTAGTATGTTCCATAGACATTTATTATATAA 208  
QY 1070 cggaaacctgaaaaaatatgct 1091  
Db 209 AACAAATGTCACACTATAGTTT 230

RESULT 4  
A0695956 521 bp DNA linear GSS 06-JUL-1999  
LOCUS HS\_2160\_A2\_P03\_T7C C1T Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=2160 Col=6 Row=K, DNA sequence.  
ACCESSION A0695956  
VERSION A0695956.1 GI:5386204  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 521)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htsc.washington.edu>  
Plate: 2160 row: K column: 6  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 521.  
Location/Qualifiers  
1..521  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate=2160 Col=6 Row=K"  
/sex="male"  
/note="Organ: Sperm; Vector: pBeloBAC11; BAC Clones in  
E-coli DH10B"

BASE COUNT 150 a 128 c 92 g 148 t 3 others  
ORIGIN

Query Match 3.5%; Score 78.4; DB 12; Length 521;  
Best Local Similarity 73.5%; Pred. No. 4.2e-09;  
Matches 100; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 906 gcaaaccttcctgtaagaacgacagacattttagcctctgycatagtc 965  
Db 290 GCAAACTTTCTGTAAAGGGCCAGATAGTAAATATGTAGACTTTCAGAGACATATGTT 349  
QY 966 ctgcagcaactactcaatctctgctctctgtagcagaagaacattagcaacaatagtc 1025  
Db 350 CTCGTTCACACTACTAGCTCTCGCGTTGTAGCACAAGCAGCCATAGACATATGCA 409  
QY 1026 aacaaactatgtgac 1041  
Db 410 GACCAATGCAATGTGAC 425

RESULT 5  
AM183569 435 bp mRNA linear EST 18-NOV-1999  
LOCUS

DEFINITION xj76h04.x1 SOARES\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2663191.3', mRNA sequence.  
ACCESSION AM183569  
VERSION AM183569.1 GI:6452083  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 435)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -400P from Glibco  
High quality sequence stop: 420.  
Location/Qualifiers  
1..435  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:2663191"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site.1: Not I; Site.2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI-CGAP-GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
1.W.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 149 a 82 c 60 g 144 t  
ORIGIN

Query Match 3.5%; Score 78.2; DB 9; Length 435;  
Best Local Similarity 59.6%; Pred. No. 4.5e-09;  
Matches 201; Conservative 0; Mismatches 108; Indels 28; Gaps 3;

QY 901 gttgggcaaaccttctctgtaagaacccgacaggaacattttaggctctgycat 960  
Db 108 GGTGAGCAAACTTTTCTGTAAAGGGCCATACATATATTTCACTCTCGACACAT 167  
QY 961 atggtctcaagtcacaactactcaatctctgctctgtagcagaagaacattagcaacaat 1020  
Db 168 TTGGTCTTTTCTGCAAACTACTCAACTCTGTATACATATACCAAAAGCTGCCACAGCAT 227  
QY 1021 atgtcaacaacaatctgagcccatgaaactatctat--tatgatacgaagacc 1077  
Db 228 ATGTAAACAATAGAGGTGACTGTAAATGTATTTATGACACTCAAAATTTGCACT 287  
QY 1078 tgaataataagcttctc-----ttgatatttccccaatcaatataaa 1121  
Db 288 TCAAAATCATTTTCATGCTTCCCAAAACATTCTTCTTTGACATTTTTCACCATTTTAA 347  
QY 1122 aacgtaaanaactctttagtcgaaggttaagccatctcaagcttagcagtgagcagggc 1181  
Db 348 AAGGTAAAAACATTTCTTAGTGTACAGGCTGTGAAAAACT-----AGTCATATAGGC 398  
QY 1182 tggatttgcttgacctacagttggccaatccctg 1218  
Db 399 TTAATCTGGCCCAAGCCATAGTGTGCCAATCCCTG 435

RESULT 6

AI125442	AI125442	459 bp	mRNA	linear	EST 28-OCT-1998
LOCUS	q689b02.x1 Soares_testis_WHT Homo sapiens cDNA clone IMAGE:173619				
DEFINITION	3', mRNA sequence.				
ACCESSION	AI125442				
VERSION	AI125442.1	GI:3593956			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 459)				
AUTHORS	NCI-CCAP <a href="http://www.ncbi.nlm.nih.gov/nciccap">http://www.ncbi.nlm.nih.gov/nciccap</a> .				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CCAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
CDNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution Information can be  
found through the I.M.A.G.E. Consortium/BLM at:  
www-bio.linn.lsu.edu/dbp/image/image.html  
Insert Length: 592 Std Error: 0.00  
Seq primer: ~40ml3 fwd. RT from Amersham.  
Location/Qualifiers  
1. .459

```

159  /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_image="IMAGE:1736619"
      /clone_lib="Soares_testis_NHT"
      /sex="male"
      /lab_host="DH10b"
      /note="Vector: pT73d-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - 01190(drr) primer 15'
TGTACCAATCTGATGAGGAGCGGCCCAATTTTCTTTT 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73s vector. Library
was through one round of normalization to Cot1, and was
constructed by Bento Soares and M. Fatima Bonaldo."
      BASE COUNT
      159 a      88 c      61 g      151 t

```

[illegible]

```
Oy 1182 cygatttgcctgtgacctacagttgccaatccctgatcccaa 1226
      ||| ||| | | | | | | | | | | | | | | | |
Db 396 TTAATCTGCCCCACAAGCCATAGTTGCCAATGCGCTGCACACTA 440
```

[illegible]

REFERENCE  
AUTHORS

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.  
1 (bases 1 to 628)  
Zhao S., Adams M.D., Niernman W., Malek J., de Jong P. and Venter J.C.

TITLE	JOURNAL	COMMENT
Use of BAC End Sequences from Library RPC1-11 for Sequence-Ready Map Building	Unpublished (1997)	
Other_GSSs: RPC11-151b24.TV		

Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: [hbee@lgr.org](mailto:hbee@lgr.org)  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong ([pieter@dejong.med.buffalo.edu](mailto:pieter@dejong.med.buffalo.edu)). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC end search page: [http://www.ligr.org/tob/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.ligr.org/tob/humgen/bac_end_search/bac_end_search.html)  
Seq primer: SP6  
Class: BAC ends.

FEATURES	
source	location/Qualifiers
1. 628	
/organism="Homo sapiens"	
/db_xref="GDB:7557647"	
/db_xref="taxon:9606"	
/clone="RPCT-11-151B24"	
/clone_11b="RPCT-11"	
/sex="Male"	
/cell_type="Lymphocytes"	
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCT11 Human Male BAC Library"	
173 a 133 c 136 g 185 t	1 others

	Query Match	3.5%;	Score 78.2;	DB 12;	Length 628;
	Best Local Similarity	72.1%;	Pred. No. 5.1e-09;		
	Matches 101;	Conservative	0;	Mismatches 39;	Indels 0; Gaps
QY	892	tgaacccaggltggcacaactcttcctgtaaaagaacacagagaactatttagtgcct	951		
Db	572	TAGGCCAGGGGTTGGCATACTTTCCTGTAAAGCACAATATGGAATATTTCATATTNT	513		
QY	952	gtatgacatatggtcagtcagcaactcatcctctgtagcagaaagcaaat	1011		
Db	512	GTCGGCCCTTATGCTGCCTCAACCAACCAATAATTCTGCAGCTGTAGCATGAACACAGCC	453		
QY	1012	agcaacatatgtccaacaaa	1031		
Db	452	ACAACACAGTAAGTAAACAAA	433		

RESULT  
AI128823



	LOCUS	A1128823	393 bp	mRNA	linear	EST 11-SEP-1996	
	DEFINITION	ga94ad06.s1 Soares_fetal_heart_NbhH19W Homo sapiens cDNA clone					
	ACCESSION	IMAGE:1694411 3', mRNA sequence.					
	VERSION	A1128823					
	KEYWORDS	A1128823.1 GI:3597337					
	SOURCE	EST.					
	ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
	REFERENCE						
	AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap.					
	TITLE	Tumor Gene Index					
	JOURNAL	Unpublished (1997)					
	COMMENT	Contact: Robert Strausberg, Ph.D. Email: rgs@phs-trmail.nhl.gov This clone is available royalty-free through LNL : contact the IMAGE Consortium (info@image.lnl.gov) for further information. Seq primer: 40mJL fwd, ET from Amerisham.					
	FEATURES	Location/Oualifiers					
	source	 . . . 393 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image:1694411" /clone_lib="Soares_fetal_heart_NbhH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" /note="Organ: heart; Vector: pPT73D (Pharmacia) with a modified polylinker; Site_1: Not I - Site_2: Eco RI; left strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAGATCTCAATGGCAGCGCCGCATCTTTTTTTTTTTT 3']; double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTV73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Felina Donaldso. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbhH19W."					
BASE COUNT	140 a	70 c	53 g	130 t			
ORIGIN							
Query Match	Best Local Similarity	3.5%	Score 77.4;	DB 9;	Length 393;		
Matches 174;	Conservative	0;	Mismatches 96;	Indels 19;	Gaps 2;		
DQ	Y	901	gtcgggcaacatcttcctctaagaagaccaggaaactatttagctcggtgccat	960			
DQ	Db	103	gggcagacaaccttttcttgtaaggcccatcacatpacatattttcagtcttcgagacacat	162			
DQ	Y	961	atggttgtcagtcacaactatcltatcttgccttcgttcagcaagaagcaattagcaaat	1020			
DQ	Db	163	ttagctgttttgtaaactaacctcaatctgcgatgaacatatcaaagaaaagctgccaacgacat	222			
DQ	Y	1021	atgtcaacaacataatgttaccccacatgaanaacttatllat---tatgatacagaacc	1077			
DQ	Db	223	atgtaaaaacaanagaggtgactgtatataaattgtattttangacactgaannttgacct	282			
DQ	Y	1078	tgaataatltgtcttct-----ttgatattttcccccaataataanaa	1121			
DQ	Db	283	tcaaanacatttttttcagtctcccaaaaacatttttacttttttgactattttttttaaccaacatttaaa	342			
DQ	Y	1122	aacgtaaaaaactactcttagtcgcaaggtlaagccatlctcagcttag	1170			
DQ	Db	343	aaagtaaaaaaaaaactttcttagtgtagacagctgtgaaaaaaactagtcataag	391			
RESULT	9						
w76494/c	w76494	411 bp	mRNA	linear	EST 17-OCT-1996		

DEFINITION	z6d7d03.r1 Soares fetal heart NBHH19W Homo sapiens cDNA clone IMAGE:345701 5', mRNA sequence.
ACTION	
ACCESSION	M76494
VERSION	E7*
KEYWORDS	WT6494.1 GI-1386739
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  1 (bases 1 to 411)
REFERENCE	Hiller,L., Clark,N., Dubique,T., Elliston,K., Hawkin,M., Holman, M., Holdman,W., Kucaba,P., Le,M., Lemmon,G., Marras,H., Parsons,J., Rifkin,L., Kohlring,T., Soares,M., Tan,F., Trevarskis,E., Watelson R., Williamson,A., Mohldmann,P. and Wilson,R. The WasNU-Merck EST Project Unpublished (1995) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wasnu.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium ( <a href="#">infoimage.lnl.gov</a> ) for further information. Insert length: 466 Std Error: 0.00 Seq primer: mob.REGA+EM
FEATURES	Location/Qualifiers 1..411 /organism="Homo sapiens" /db_xref="GDB:1271076" /db_xref="taxon:9606" /clone="IMAGE:345701" /clone_1lb-"Soares_fetal_heart_NBHH19W" /sev="unknown" /gen_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" /note="Organ: heart; Vector: pRTV3D (Pharmacia) with a modified polylinker; Site_1: Not I ; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5' TGTTCACAATCGAAAGTGCGAGCAGCCGCCTATTCTTTTGGTTTTTGT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTV3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Felipe Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
BASE COUNT	129 A     58 C     79 G     145 T
ORIGIN	NBHL19W,"
Query Match	3.5%; Score 77.4; DB 10; Length 411;
Best Local Similarity	60.2%: Fred.No.:7.le-09; Indels 19; Gaps 2;
Matches 174; Conservative	0; Mismatches 96;
Dy	325 GGTCAGCAACTTTTCTGTATAAGGCCCATTAACATATTTTCAGCTCTGCAGACCMT 266
OY	961 atgtctcagtcacaaacttcatatcttgccctcgtagcaagaagaattgacaacct 1020
Dy	265 TTGGTCTTTTCTGAACACTCACCTGCTGTACATTAACAAAAGCTGCCACAGCAAT 206
OY	1021 atgtcaaacaatalgatgcaccatgaanaacttatattal--talgalatagaaac 1077
Dy	205 ATGTAAAAAATGAGSGGTGATGTAAATGTATTTTAGAACACTGAAATTTGAOCT 146
OY	1078 tgaataaatagtcttct-----tttgtatttcccacaatatnaaaa 1121
Dy	145 TCATATCATTTTCATCTCTCCCCAANAACATTCTACTTTTGACTATTTTTCACACATTAAA 86

```
OY 1122 aacgtaaaaactactctcttagtcgcaaggttaagccattctcagcttag 1170
      || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 AAAGTAAACCATTTCTTAGGTGACAGGCTGTAGAAAACTAGTCATAG 37
```









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 05:33:21 ; Search time 485.09 Seconds

(Without alignments)  
7928.195 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240

Perfect score: 2240

Sequence: 1 taatcagactcaactatagg9.....aaatccacaagaacctgc 2240

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	100.0	2286	22	AAH26595
2	81.8	3.7	37437	22	AAK84165
3	81.4	3.6	37442	22	AAK84166
4	81	3.6	5294	22	AAH14600
5	77.6	3.5	26591	22	AAI36313
6	75.8	3.4	481	22	AAK74761
7	73	3.3	1937	22	AAH04430
8	71.6	3.2	628	22	AAH98619
9	71.4	3.2	286	21	AAA41353

C 10	71.2	3.2	407	21	AAC27788	Human secreted pro
C 11	70	3.1	123219	23	AAH88703	Human DNA sequence
C 12	69.8	3.1	50000	20	AAH23517	Human kidney amino
C 13	69.4	3.1	8029	22	AAH9747	Human protein enco
C 14	69.4	3.1	46765	24	AAS99306	DNA encoding Aldeh
C 15	69.4	3.1	47319	22	AAK64813	Human immune/haema
C 16	69.4	3.1	47319	22	AAK72230	Human immune/haema
C 17	69.4	3.1	236303	22	AAH11614	Human genomic DNA
C 18	68.6	3.1	325791	22	AAH43104	Human Oestrogen re
C 19	68.2	3.0	97662	22	AAH83908	Genomic sequence o
C 20	68	3.0	31140	22	ABA08065	Human ovarian and
C 21	68	3.0	31140	22	AAH06791	Human reproductive
C 22	67.8	3.0	17700	22	AAK68945	Human immune/haema
C 23	67.6	3.0	2515	22	AAH17444	Human granulocyte
C 24	67.6	3.0	6235	21	AAH29169	Human G-CSF genom
C 25	67.6	3.0	6679	21	AAH29170	Human granulocyte
C 26	67.6	3.0	6679	22	AAH17443	Human granulocyte
C 27	66.6	2.9	289	22	AAS39009	Novel human diagno
C 28	66.4	3.0	466	22	AAH33971	Human colon cancer
C 29	66.4	3.0	655	22	AAH17197	Human breast cance
C 30	66	2.9	871	22	AAH04280	Human CDNA clone (
C 31	66	2.9	1712	22	AAH13903	Human CDNA sequenc
C 32	66	2.9	2291	22	AAH17522	Human CDNA sequenc
C 33	66	2.9	7496	22	AAK70490	Human immune/haema
C 34	66	2.9	7496	22	AAK74614	Human immune/haema
C 35	66	2.9	10225	22	AAH99021	Human kidney relat
C 36	66	2.9	10225	22	AAH99021	Human kidney relat
C 37	66	2.9	24292	22	AAH15839	Human nervous syst
C 38	66	2.9	28731	22	AAK64887	Human immune/haema
C 39	66	2.9	28731	22	AAK69666	Human immune/haema
C 40	66	2.9	28731	22	AAK70491	Human immune/haema
C 41	66	2.9	28731	22	AAK70491	Human immune/haema
C 42	66	2.9	28731	22	AAK84942	Human immune/haema
C 43	65.6	2.9	872	22	AAH98871	Human immune/haema
C 44	65.4	2.9	519	22	AAH35469	Human EST-derived
C 45	64.8	2.9	15044	22	AAS36290	Human musculoskele
						Human cardiovascular

## ALIGNMENTS

RESULT 1	
AAH26595	standard; DNA; 2286 BP.
AAH26595;	
12-NOV-2001 (first entry)	
Human melanoma differentiation associated gene-7 (Mda-7) promoter.	
Melanoma differentiation associated gene-7; Mda-7; promoter; human;	
neuroblastoma; astrocytoma; glioblastoma multiforme;	
cervical cancer; breast cancer; colon cancer; prostate cancer;	
osteosarcoma; chondrosarcoma; tumour; therapy; ds.	
Homo sapiens.	
Key	Location/Qualifiers
promoter	1..2240
protein_bind	/tag= a
protein_bind	138..146
protein_bind	/tag= b
protein_bind	/bound_moiety= "C/EBP"
protein_bind	446..454
protein_bind	/tag= c
protein_bind	/bound_moiety= "C/EBP"
protein_bind	773..777
protein_bind	/tag= d
protein_bind	/bound_moiety= "C/EBP"
protein_bind	801..807
protein_bind	/tag= e
protein_bind	/bound_moiety= "AP-1"

```

FT      protein_bind      1109..1113
FT      /*tag= f
FT      /bound_molecety= "C/EBP"
FT      protein_bind      1207..1214
FT      /*tag= g
FT      /bound_molecety= "AP1"
FT      protein_bind      1209..1214
FT      /*tag= h
FT      /bound_molecety= "C/EBP"
FT      protein_bind      1530..1537
FT      /*tag= i
FT      /bound_molecety= "AP1"
FT      protein_bind      2115..2121
FT      /*tag= j
FT      /bound_molecety= "AP1"
FT      protein_bind      2126..2133
FT      /*tag= k
FT      /bound_molecety= "C/EBP"
FT      TATA_signal      2210..2214
FT      /*tag= l
FT
FT      WO200164921-A1.
FT
FT      07-SEP-2001.
FT
FT      28-FEB-2001; 2001WO-US06782.
FT
FT      29-FEB-2000; 2000US-0515369.
FT
FT      (UYCO ) UNIV COLUMBIA NEW YORK.
FT
FT      Fisher PB, Madiredi MT:
FT      WPI: 2001-565508/63.
FT
FT      Melanoma differentiation associated gene-7 promoter capable of
FT      treating cancer comprises directing transcription of heterologous
FT      coding sequence encoding tumor suppressor polypeptide positioned
FT      downstream, useful for treating cancer
FT
FT      Claim 1: Fig 6A-B; 132pp; English.
FT
XX      CC      The present sequence is that of the promoter region of the human
XX      CC      melanoma differentiation associated gene-7 (Mda-7). It includes
XX      CC      2240 bp from the 5' flanking region of the Mda-7 gene. The
XX      CC      promoter was isolated from a human placental genomic library
XX      CC      using a PCR-based method. The Mda-7 promoter exhibits melanocyte
XX      CC      tissue specificity and can only be activated in the targeted
XX      CC      tissue, i.e. the skin. Therefore, a gene of interest driven by
XX      CC      the Mda-7 promoter will be differentially expressed in these cells,
XX      CC      minimizing systemic toxicity. A recombinant expression construct
XX      CC      in which the human Mda-7 promoter is operably linked to a coding
XX      CC      sequence encoding a tumor suppressor protein is claimed. The
XX      CC      tumor suppressor is preferably p21, retinoblastoma protein or p53.
XX      CC      A host cell comprising the expression construct is also claimed,
XX      CC      and is preferably a tumor cell selected from a melanoma,
XX      CC      neuroblastoma, astrocytoma, glioblastoma multiforme, cervical
XX      CC      cancer, breast cancer, lung cancer or prostate cancer cell.
XX      CC      The promoter preferably comprises nucleotides 1-2240 of the present
XX      CC      sequence. A pharmaceutical composition including the recombinant
XX      CC      expression construct is used in a claimed method of treating
XX      CC      melanoma, neuroblastoma, astrocytoma, glioblastoma multiforme,
XX      CC      cervical cancer, breast cancer, colon cancer, prostate cancer,
XX      CC      osteosarcoma, chondrosarcoma or a cancer of the central nervous
XX      CC      system.
XX
XX      Sequence 2286 BP; 547 A; 592 C; 510 G; 637 T; 0 other:

```

```

Query Match      100.0%; Score 2240; DB 22; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1      taatagactactatagggcgctgactgatcaccttttgaacccaggtctgcgcct 60
QY      |||||||
DB      1      taatacgaactacacataagggcgctgactgatcaccttttgaacccaggtctgcgcct 60
QY      61      ccaaaagcttgactacatactagatctcaactgtagtctggcgaaggttcctagttct 120
QY      |||||||
DB      61      ccaaaagcttgactacatactagatctcaactgtagtctggcgaaggttcctagttct 120
QY      121      ctcttgacctctctcttgagatgaatagtctatgaataagctcatcgagagcgagcc 180
QY      |||||||
DB      121      ctcttgacctctctcttgagatgaatagtctatgaataagctcatcgagagcgagcc 180
QY      181      aggcacatgttgcctggaactatccatgltatagatctctcccgagacaggtagct 240
QY      |||||||
DB      181      aggcacatgttgcctggaactatccatgltatagatctctcccgagacaggtagct 240
QY      241      actacagatcccaaggtgtacccctgagggcgaaggtgtatccatgactcctgcct 300
QY      |||||||
DB      241      actacagatcccaaggtgtacccctgagggcgaaggtgtatccatgactcctgcct 300
QY      301      gtccagcctgaccttaacagatccatccacccctgcctgcctcccgctatctgcaga 360
QY      |||||||
DB      301      gtccagcctgaccttaacagatccatccacccctgcctgcctcccgctatctgcaga 360
QY      361      cagtagtctagatltcaagctgcctgggggtcatttccctcagctcctgcctt 420
QY      |||||||
DB      361      cagtagtctagatltcaagctgcctgggggtcatttccctcagctcctgcctt 420
QY      421      gctgtctcctgcctcccaactatcaactacacagacactccacactgcctctctctg 480
QY      |||||||
DB      421      gctgtctcctgcctcccaactatcaactacacagacactccacactgcctctctctg 480
QY      481      tctcatcactgctctctgacatcttlatcctatagtagttagtgaggttcttgfaac 540
QY      |||||||
DB      481      tctcatcactgctctctgacatcttlatcctatagtagttagtgaggttcttgfaac 540
QY      541      gccctaaatccacatggtgggaagggaggttgagggaaggtgctgtgaggtcgt 600
QY      |||||||
DB      541      gccctaaatccacatggtgggaagggaggttgagggaaggtgctgtgaggtcgt 600
QY      601      gctactctctgagggtagaactcgggacctccagaaacaaaggtatcaagctgtgca 660
QY      |||||||
DB      601      gctactctctgagggtagaactcgggacctccagaaacaaaggtatcaagctgtgca 660
QY      661      gctatagccaagcagactcgtcggcgaaggtatgcaaaaggtattgttgcttaaga 720
QY      |||||||
DB      661      gctatagccaagcagactcgtcggcgaaggtatgcaaaaggtattgttgcttaaga 720
QY      721      aataaacacactgagtagatgagatggaggggtgtgtgtgcgaagagattgggaa 780
QY      |||||||
DB      721      aataaacacactgagtagatgagatggaggggtgtgtgtgcgaagagattgggaa 780
QY      781      gagctgcgaaggtgtgttctactactctctctcttcttcaactcagctagctgtg 840
QY      |||||||
DB      781      gagctgcgaaggtgtgttctactactctctctcttcttcaactcagctagctgtg 840
QY      841      aggcagttatcctgtgccccaagctacatctccactccgttcccaagctggaacag 900
QY      |||||||
DB      841      aggcagttatcctgtgccccaagctacatctccactccgttcccaagctggaacag 900
QY      901      gttgggcaaacctctcctgttaaagaacagacaggaactatttagcctgtgccaat 960
QY      |||||||
DB      901      gttgggcaaacctctcctgttaaagaacagacaggaactatttagcctgtgccaat 960
QY      961      atgtctcagttacacactactatctctgcctctgtagcagaaagcatttagacaacat 1020
QY      |||||||
DB      961      atgtctcagttacacactactatctctgcctctgtagcagaaagcatttagacaacat 1020
QY      1021      atgttaaacataatgtgaccccatgaaactattattatgatacaggaacactgga 1080
QY      |||||||
DB      1021      atgttaaacataatgtgaccccatgaaactattattatgatacaggaacactgga 1080
QY      1081      aataaatgtcttctttagatttctcccaatcatlaaaaaagctaaaaactactctta 1140

```



|||||  
Db 1081 aaataagctcttcttgatttttccccaatcataaataaactactctta 1140  
Qy 1141 ggtcgcaaggttaagcattcttcagcttagagtgagcagcttgattgcttgacct 1200  
Db 1141 ggtcgcaaggttaagcattcttcagcttagagtgagcagcttgattgcttgacct 1200  
Qy 1201 acaattgscatccctcgatctcccaaatgtatctcccaaggagtgaggcaataactat 1260  
Db 1201 acaattgscatccctcgatctcccaaatgtatctcccaaggagtgaggcaataactat 1260  
Qy 1261 ggggaagtgcctggaatcaacagagtttaagagcatcagacattccagagcgctagcac 1320  
Db 1261 ggggaagtgcctggaatcaacagagtttaagagcatcagacattccagagcgctagcac 1320  
Qy 1321 atgcagaaggctcttaactaacctcattgattatctgtttcaaggaaacttgaaag 1380  
Db 1321 atgcagaaggctcttaactaacctcattgattatctgtttcaaggaaacttgaaag 1380  
Qy 1381 acaagaattcccaaacctagagtcagagagctgctcttgagaaacacgtctctgctg 1440  
Db 1381 acaagaattcccaaacctagagtcagagagctgctcttgagaaacacgtctctgctg 1440  
Qy 1441 atgcacctcactgggcacatgtagaatctagagctgagctgctctgctagcttgagatag 1500  
Db 1441 atgcacctcactgggcacatgtagaatctagagctgagctgctctgctagcttgagatag 1500  
Qy 1501 gtccagagctcttgactgctccctgagctgagacacacacgtctgctgctccctgagt 1560  
Db 1501 gtccagagctcttgactgctccctgagctgagacacacacgtctgctgctccctgagt 1560  
Qy 1561 ggttcacaagccacacagagcccaagaactagcccacacagacagcgctcccaactctctg 1620  
Db 1561 ggttcacaagccacacagagcccaagaactagcccacacagacagcgctcccaactctctg 1620  
Qy 1621 ggtctgcacatgctgatacttctcttctctgtgttccagaacctagagagagcgac 1680  
Db 1621 ggtctgcacatgctgatacttctcttctctgtgttccagaacctagagagagcgac 1680  
Qy 1681 atttggttgaagttatataacctagagacacatggttagctgcatgctcgagaaacatc 1740  
Db 1681 atttggttgaagttatataacctagagacacatggttagctgcatgctcgagaaacatc 1740  
Qy 1741 ctcaactctctgcccctgagatlaaagagaggtactacaggtattcttctctgctg 1800  
Db 1741 ctcaactctctgcccctgagatlaaagagaggtactacaggtattcttctctgctg 1800  
Qy 1801 ggaactactgcccagcagatagatagggacatcatatgaattatttgacagagagagac 1860  
Db 1801 ggaactactgcccagcagatagatagggacatcatatgaattatttgacagagagagac 1860  
Qy 1861 tgggtgtatgctgcacagtaataatgtaattacatggtgacagagttaccagaacacctct 1920  
Db 1861 tgggtgtatgctgcacagtaataatgtaattacatggtgacagagttaccagaacacctct 1920  
Qy 1921 ggtgtgttttgccttgctttatatacattgggacaaattttaaattatatacagcg 1980  
Db 1921 ggtgtgttttgccttgctttatatacattgggacaaattttaaattatatacagcg 1980  
Qy 1981 agactcgagcgacagagaaagctaaagagactgcccctgcccacacagcgagtgtaagcc 2040  
Db 1981 agactcgagcgacagagaaagctaaagagactgcccctgcccacacagcgagtgtaagcc 2040  
Qy 2041 tgaactccaacccaggtctcatctacacctcagagggctgttcccatcgctgattgctc 2100  
Db 2041 tgaactccaacccaggtctcatctacacctcagagggctgttcccatcgctgattgctc 2100  
Qy 2101 cttaaaagtgaagggtactagagcaatgaagtaattctcagagaaacatgacaaatttcc 2160  
Db 2101 cttaaaagtgaagggtactagagcaatgaagtaattctcagagaaacatgacaaatttcc 2160  
Qy 2161 ctcttcacactcctcttcttctccacccctccccaatcaacccccaatatatgccc 2220  
Db 2161 ctcttcacactcctcttcttctccacccctccccaatcaacccccaatatatgccc 2220

Db 2161 ctcttcacactcctcttcttctccacccctccccaatcaacccccaatatatgccc 2220  
Qy 2221 aaatcccaacaagccttgc 2240  
Db 2221 aaatcccaacaagccttgc 2240  
RESULT 2  
AAK84165/c  
ID AAK84165 standard; DNA; 37437 BP.  
XX  
AC AAK84165;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38977.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytosolic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX  
PN MO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.



RESULT 3  
AAK84166/c  
ID AAK84166 standard; DNA: 37442 BP.  
XX  
AC AAK84166;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human Immune/haematopoietic antigen genomic sequence SFO ID NO:38978.  
XX  
KM Human: Immune; haematopoietic; Immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189674.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214866.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.



DB 1795 AGACATATGTACGCA 1779

RESULT 5  
AAL36313/c  
ID AAL36313 standard; DNA; 26591 BP.  
XX  
AC AAL36313;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.  
XX  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilicer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US01338.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0234999.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.











DB 224 ACATTGAATTGAA 210

## RESULT 11

AAH8703  
ID AAH8703 standard; DNA; 123219 BP.  
XX  
XX  
AC AAH8703;  
XX  
XX 26-FEB-2002 (first entry)  
XX  
XX  
DE Human DNA sequence SEQ ID 543.  
XX  
XX Single nucleotide polymorphism; SNP; diallelic marker; human;  
XX  
XX central nervous system disorder; CNS; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200151659-A2.  
XX  
XX 19-JUL-2001.  
XX  
XX 11-JAN-2001; 2001WO-1B00116.  
XX  
XX 13-JAN-2000; 2000US-0175854.  
XX  
XX (GEST ) GENSET.  
XX  
XX Chu T, Blumenfeld M, Cohen D;  
XX  
XX WPI; 2001-483085/52.  
XX  
XX Isolated polynucleotides, useful for genotyping nucleic acids for  
XX  
XX diallelic markers for the diagnosis of depression, comprises central  
XX  
XX nervous system disorder related diallelic marker -  
XX  
XX disclosure; Page 439-472; 51pp; English.  
XX  
XX The present invention relates to diallelic markers derived from human  
XX  
XX genes involved in central nervous system (CNS) disorders (see  
XX  
XX AAH88161-AAH88702). The markers have a single nucleotide polymorphism  
XX  
XX (SNP) and are useful in determining the genetic predisposition of  
XX  
XX individuals to CNS disorders, by identifying the nucleotides at a set of  
XX  
XX genetic markers in a biological sample, where the markers comprise at  
XX  
XX least one CNS disorder related marker. The present sequence was used  
XX  
XX to illustrate the invention.  
XX  
XX Sequence 123219 BP; 27921 A; 32327 C; 32369 G; 30437 T; 165 other;

Query Match 3.1%; Score 70; DB 23; Length 123219;  
Best Local Similarity 61.5%; Pred. No. 9.7e-09;  
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 895 acccaggttgaggcaaacctctctgttaagaacacagacagacatttttagctctgtc 954  
DB 17375 agcagggggttcgaaccttcttcggaagggatagatgaacgtgcataggcttctg 17434  
QY 955 tgcacatagctcagtcacaaactactctctgcctcgttagcagaagaactaac 1014  
DB 17435 ggcacatggtcctcatcgacatcactctatgctgttagcagaagaagcaca 17494  
QY 1015 aacacatactgcaaacacatactgacccacatgaactattttatgatacagaa 1074  
DB 17495 gacacatagtcacaagaagcgctgctgcatgcacataaactatttataaacaac 17554  
QY 1075 ac 1076  
DB 17555 ac 17556

RESULT 12

AAK3517/C  
ID AAK3517 standard; DNA; 50000 BP.  
XX  
XX  
AC AAK3517;  
XX  
XX 23-JUN-1999 (first entry)  
XX  
XX  
DE Human kidney aminopeptidase P genomic DNA fragment 1.  
XX  
XX  
XX Aminopeptidase; human; Amp; gene therapy; treatment; Amp-deficiency;  
XX  
XX prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;  
XX  
XX arterial stenosis; industrial protein feed; malabsorption syndrome;  
XX  
XX proteinaceous waste degradation; additive; immunohistochemistry; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO9911799-A2.  
XX  
XX 11-MAR-1999.  
XX  
XX 02-SEP-1998; 98WO-US18426.  
XX  
XX 02-SEP-1997; 97US-0057854.  
XX  
XX (MEDT-) MEDICAL COLLEGE GEORGIA RES INST.  
XX  
XX Ryan JW, Sprinkle JJC, Venema RC;  
XX  
XX WPI; 1999-205193/17.  
XX  
XX Nucleic acid encoding human aminopeptidase P  
XX  
XX  
XX Claim 13; Page 80-109; 201pp; English.

This invention describes the isolation of a novel human aminopeptidase P (Amp). This protein is used to produce recombinant Amp and can be used for gene therapy for treating Amp-deficiency conditions. Its fragments are used as primers and probes to identify patients with homozygous and heterozygous Amp deficiency, including prenatal diagnosis (patients defective in Amp are at risk of developing angioedema if treated with angiotensin-converting enzyme inhibitors), also as antisense inhibitors in cases of excessive Amp expression. The product of the invention is also used to identify Amp-expressing sequences in other animals and to generate transgenic animals, and comparisons of genomic sequences are used to detect mutations. Amp inhibitors are potentially useful as antihypertensive agents and to prevent or treat arterial (re)stenosis or atherosclerosis. The structure of Amp is used to design synthetic substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal imido bonds, can be used to degrade industrial protein feeds to free amino acids, to degrade proteinaceous wastes, as additives in enzyme formulations used to treat malabsorption syndrome and for studying its biological role. Antibodies against Amp are used in immunohistochemical methods to study Amp distribution.

Sequence 50000 BP; 13187 A; 12125 C; 11842 G; 12846 T; 0 other;

Query Match 3.1%; Score 69.8; DB 20; Length 50000;  
Best Local Similarity 72.7%; Pred. No. 6.7e-09;  
Matches 104; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 889 gcttgaccacaggttgaggcaaacctctctgttaagaacacagacagacatttttagc 948  
DB 8967 gcttgaccacaggttgaggcaaacctctctgttgaaggctcagatgtaagatttttagc 8910  
QY 949 tctgtgtgcacatagctcagtcacaaactactcatctcgtctgttagcagaagaag 1008  
DB 8909 tcttcgacacacatgacgtctctgtacacacacactcaactcgtctgttagcagaagaag 8850  
QY 1009 atgacacacacatgacacacaa 1031  
DB 8849 actatagacacacatgacacacaa 8827



```

FT      /*tag= h
FT      /product= "ALDH5A1"
FT      /note= "Aldehyde dehydrogenase 5 family member A1.
FT      Specifically claimed in claim 25"
FT      replace(4059,G)
FT      /*tag= i
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4111,A)
FT      /*tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4155,C)
FT      /*tag= k
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4179,G)
FT      /*tag= l
FT      /standard_name= "Single nucleotide polymorphism"
FT      4404..11575
FT      /*tag= m
FT      /number= 1
FT      replace(4430,G)
FT      /*tag= n
FT      /standard_name= "Single nucleotide polymorphism"
FT      11576..11659
FT      /*tag= o
FT      /number= 2
FT      11660..12315
FT      /*tag= p
FT      /number= 2
FT      12316..12486
FT      /*tag= q
FT      /number= 3
FT      replace(12415,T)
FT      /*tag= r
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(12422,T)
FT      /*tag= s
FT      /standard_name= "Single nucleotide polymorphism"
FT      12487..13921
FT      /*tag= t
FT      /number= 3
FT      replace(12559,T)
FT      /*tag= u
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(13855,G)
FT      /*tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(13917,T)
FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      13922..14038
FT      /*tag= x
FT      /number= 4
FT      replace(14021,T)
FT      /*tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      14039..24219
FT      /*tag= z
FT      /number= 2
FT      replace(14046,A)
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(14179,A)
FT      /*tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      24220..24363
FT      /*tag= ac
FT      /number= 5
FT      24364..29453
FT      /*tag= ad
FT      /number= 5
FT      24598..31819
FT      /*tag= ae
FT      /number= 6

```

```

FT      variation
FT      /*tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      29454..29597
FT      /*tag= ag
FT      /number= 6
FT      replace(29469,A)
FT      /*tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      31820..31978
FT      /*tag= ai
FT      /number= 7
FT      31979..37049
FT      /*tag= aj
FT      /number= 7
FT      37050..37219
FT      /*tag= ak
FT      /number= 8
FT      37220..41171
FT      /*tag= al
FT      41172..41230
FT      /*tag= am
FT      /number= 9
FT      replace(41217,C)
FT      /*tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      41231..42559
FT      /*tag= ao
FT      /number= 9
FT      replace(41237,A)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(41403,C)
FT      /*tag= aq
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(42380,T)
FT      /*tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      42560..42765
FT      /*tag= as
FT      /number= 10
FT      WO200190119-A2.
FT      PN
FT      29-NOV-2001.
FT      PD
FT      XX
FT      21-MAY-2001; 2001WO-US16558.
FT      PF
FT      XX
FT      19-MAY-2000; 2000US-205849P.
FT      PR
FT      XX
FT      (GENA-) GENMAISSANCE PHARM INC.
FT      PA
FT      XX
FT      PI
FT      KIlem SE, Koshy B, Tanguay DA;
FT      WPI. 2002-089912/12.
FT      DR
FT      P-PSDB; AAU73594.
FT      XX
FT      New genetic variants of human aldehyde dehydrogenase 5 family, member
FT      AL, ALDH5A1 gene for treating metabolic diseases and for expressing
FT      ALDH5A1 protein useful in identifying drugs to treat 4-hydroxybutyric
FT      aciduria -
FT      XX
FT      Disclosure; Ptg 1; 151pp; English.
FT      PS
FT      XX
FT      The invention describes an isolated polynucleotide comprising a
FT      nucleotide sequence which is a polymorphic variant of a reference
FT      sequence for the aldehyde dehydrogenase 5 family, member A1
FT      CC (succinate-semialdehyde dehydrogenase) (ALDH5A1) gene or its fragment.
FT      CC The polypeptide is useful for screening for drugs targeting it by
FT      CC contacting the ALDH5A1 polymorphic variant with a candidate agent and
FT      CC assaying for binding activity. The polypeptide and haplotypes are useful
FT      CC for identifying an association between a trait such as a clinical

```

CC response to a drug targeting ALDH5A1 and a haplotype ALDH5A1 gene.  
CC Transgenic animals are also useful for studying expression of the ALDH5A1  
CC isogenes in vivo, for in vivo screening and testing of drugs against  
CC ALDH5A1 protein and for testing the efficacy of therapeutic agents and  
CC compounds for 4-hydroxybutyric aciduria and metabolic diseases in a  
CC biological system. Antibodies are useful for diagnostic and prognostic  
CC formats and therapeutic methods, for immunoprecipitating the polypeptide  
CC from solution, for detecting ALDH5A1 protein isoforms in biological  
CC samples, frozen tissue sections, for use in immunocytochemical,  
CC immunohistochemical and immunofluorescence techniques. The polynucleotide  
CC is useful for gene therapy and antisense gene therapy. This sequence  
CC encodes ALDH5A1 (located on chromosome 6p22), and forms a reference  
CC sequence on which polymorphic sequences encoding isoforms of the ALDH5A1  
CC protein are based, described in the method of the invention.

Query Match 3.1%; Score 69.4; DB 24; Length 46765;  
Best Local Similarity 57.2%; Pred. No. 8.4e-09;  
Matches 202; Conservative 0; Mismatches 121; Indels 30; Gaps 3;

QY 889 gctctgagcccaaggttggcacaactctctgttaagaccagacgaactatttaggc 948  
DB 45589 GTCTAGATCAGAGAGCTGGCAAACTTTTCTTTAGAGACAAGTAATCTTTTAGGC 45530  
QY 949 tctgtgtgcataatgtctcgaagcaactcattctctgtcctctgtagcac----- 1001  
DB 45529 TTTTCAGGCTATCTGCTCCCTGTCCCACTATTCACCTGTGCCAAGAACAGTCCGAGA 45470  
QY 1002 -----gaagcaattagcaacaatattgtcaacaataatgaccccatga 1048  
DB 45469 CAATACCAATGAATGAGGTGGCTATGCTCTAATATACATTTTGTATTGACACATGA 45410  
QY 1049 aaacttattt-attatgatalacggaacctgaaataatgtctcttctttagatttttc 1107  
DB 45409 AATCTGAATTTTCATAGAAATTTTTCAGTGTGATGAATATTTGTCTTTTGTGATTTCTG 45350  
QY 1108 cccaatcaataaaaaacgtaaaaactctcttaggttcgaaggttaagccattctcagct 1167  
DB 45349 CCCAATCATTTTGAATAATATAACCATTTCTTAGCTCACAGG-----TTGTATATAA 45299  
QY 1168 tagcagtcgacgagctgattgtgttgacctacagttgacccatcccat 1220  
DB 45298 AACAGGTGGCAGGCCAGGCTTTGATTGCTGGGCCATAGTTTGCACACCCTGAT 45246

RESULT 15  
AAK64813/c  
ID AAK64813 standard; DNA: 47319 BP.  
XX  
AC AAK64813;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19625.  
XX  
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229387.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235634.  
PR 27-SEP-2000; 2000US-0235636.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.

20-OCT-2000: 2000US-0241221.  
 PR 20-OCT-2000: 2000US-0241785.  
 PR 20-OCT-2000: 2000US-0241786.  
 PR 20-OCT-2000: 2000US-0241787.  
 PR 20-OCT-2000: 2000US-0241808.  
 PR 20-OCT-2000: 2000US-0241809.  
 PR 20-OCT-2000: 2000US-0241826.  
 PR 01-NOV-2000: 2000US-0244617.  
 PR 08-NOV-2000: 2000US-0246474.  
 PR 08-NOV-2000: 2000US-0246475.  
 PR 08-NOV-2000: 2000US-0246476.  
 PR 08-NOV-2000: 2000US-0246477.  
 PR 08-NOV-2000: 2000US-0246478.  
 PR 08-NOV-2000: 2000US-0246523.  
 PR 08-NOV-2000: 2000US-0246524.  
 PR 08-NOV-2000: 2000US-0246525.  
 PR 08-NOV-2000: 2000US-0246526.  
 PR 08-NOV-2000: 2000US-0246527.  
 PR 08-NOV-2000: 2000US-0246528.  
 PR 08-NOV-2000: 2000US-0246532.  
 PR 08-NOV-2000: 2000US-0246609.  
 PR 08-NOV-2000: 2000US-0246610.  
 PR 08-NOV-2000: 2000US-0246611.  
 PR 08-NOV-2000: 2000US-0246613.  
 PR 17-NOV-2000: 2000US-0249207.  
 PR 17-NOV-2000: 2000US-0249208.  
 PR 17-NOV-2000: 2000US-0249209.  
 PR 17-NOV-2000: 2000US-0249210.  
 PR 17-NOV-2000: 2000US-0249211.  
 PR 17-NOV-2000: 2000US-0249212.  
 PR 17-NOV-2000: 2000US-0249213.  
 PR 17-NOV-2000: 2000US-0249214.  
 PR 17-NOV-2000: 2000US-0249215.  
 PR 17-NOV-2000: 2000US-0249216.  
 PR 17-NOV-2000: 2000US-0249217.  
 PR 17-NOV-2000: 2000US-0249218.  
 PR 17-NOV-2000: 2000US-0249244.  
 PR 17-NOV-2000: 2000US-0249245.  
 PR 17-NOV-2000: 2000US-0249264.  
 PR 17-NOV-2000: 2000US-0249265.  
 PR 17-NOV-2000: 2000US-0249297.  
 PR 17-NOV-2000: 2000US-0249299.  
 PR 17-NOV-2000: 2000US-0249300.  
 PR 01-DEC-2000: 2000US-0250160.  
 PR 01-DEC-2000: 2000US-0250391.  
 PR 05-DEC-2000: 2000US-0251030.  
 PR 05-DEC-2000: 2000US-0251988.  
 PR 05-DEC-2000: 2000US-0256719.  
 PR 06-DEC-2000: 2000US-0251479.  
 PR 08-DEC-2000: 2000US-0251856.  
 PR 08-DEC-2000: 2000US-0251868.  
 PR 08-DEC-2000: 2000US-0251869.  
 PR 08-DEC-2000: 2000US-0251989.  
 PR 08-DEC-2000: 2000US-0251990.  
 PR 11-DEC-2000: 2000US-0254097.  
 PR 05-JAN-2001: 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Disclosure: SEQ ID NO 19625; 3071bp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytosstatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK7694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK51942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 47319 BP; 13708 A; 8453 C; 9720 G; 15438 T; 0 other;

Query Match 3.1%; Score 69.4; DB 22; Length 47319;  
 Best Local Similarity 66.2%; Pred. No. 8.5e-09;  
 Matches 100; Conservative 0; Mismatches 51; Indels 0; Gaps 0;  
 QY 881 ttcccatgctcgtgacccaggttggcaaacctctctgttaagaacacaggaacta 940  
 Db 18549 TTTACTAGCTCTAAATCAGAGATTGGCAACTTTTCTGTCAAGCCGCAATTAATTA 18490  
 QY 941 tttaggcctcgtgctgcataatgctcagtcacacactcactcctcgtcgtagca 1000  
 Db 18489 TTTTAGCCTTTTGTAAAGCCACACACAGCTCTACGGCAGCACCACTCTGTGCATTAATGCC 18430  
 QY 1001 cgaagcaattagcaacaatcgtcacaacaa 1031  
 Db 18429 CCNAAGCTGCTACAGATGATATATTAATGAA 18399

Search completed: August 16, 2002, 05:43:35  
 Job time: 11519 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:40:06 ; Search time 3499.63 Seconds

(without alignments)  
8816.372 Million cell updates/sec

Title: US-09-515-369B-1

Perfect score: 2286  
Sequence: 1 taatacgcctacataggg.....tgacttcacgcctcggagc 2286

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estbda:\*
- 2: em\_estbhm:\*
- 3: em\_estln:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hlc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hlc:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	3.5	480	10 B1262402	B1262402 602953836
2	80.6	3.5	354	10 T58770	T58770 yb80h02.s1
3	80.4	3.5	390	10 L44398	L44398 HUMEST1E3 H
4	78.4	3.4	521	12 A0695956	A0695956 HS-2160_A
5	78.2	3.4	435	9 AM183569	AM183569 xj76h04.x
6	78.2	3.4	459	9 A1125442	A1125442 qd89b02.x
7	78.2	3.4	628	12 A0378792	A0378792 RPI11-15
8	77.4	3.4	393	9 A1128823	A1128823 qe94d06.s
9	77.4	3.4	411	10 W76494	W76494 zd67d03.r1
10	77.4	3.4	680	12 AG154552	AG154552 Pan trogl
11	77.2	3.4	400	9 A1926274	A1926274 w043g06.x
12	77	3.4	322	9 A1240516	A1240516 qh12h02.x
13	77	3.4	386	9 AA443938	AA443938 zv51c03.s
14	77	3.4	393	9 AA444117	AA444117 zv51c03.r
15	77	3.4	417	9 A1208768	A1208768 q938c10.x
16	76.8	3.4	509	10 B1061153	B1061153 t13-UT011
17	76.6	3.4	327	12 B98848	B98848 CIT-HSP-228

18	76.2	3.3	499	12 A0277351	A0277351 CITBI-EI-
19	76.2	3.3	644	12 A0039307	A0039307 CIT-HSP-2
20	76.2	3.3	5529	12 AF101969	AF101969 AF101969
21	76	3.3	573	12 A0633901	A0633901 RPCI-11-4
22	75.6	3.3	530	12 A0508112	A0508112 RPCI-11-2
23	75.6	3.3	628	12 A0020407	A0020407 CIT-HSP-2
24	75.6	3.3	799	10 BG535645	BG535645 60256353
25	75.2	3.3	488	9 AV604853	AV604853 AV604853
26	74.6	3.3	348	12 A0035111	A0035111 CIT-HSP-2
27	73.6	3.2	640	12 AG093736	AG093736 Pan trogl
28	73.2	3.2	381	9 AM574882	AM574882 UT-HF-BN0
29	73.2	3.2	512	12 A0768179	A0768179 HS-3160_B
30	73	3.2	519	12 A0029411	A0029411 RPCI11-39
31	72.8	3.2	302	9 A1559884	A1559884 hb7a12.x
32	72.8	3.2	427	12 A0815705	A0815705 HS-5394_A
33	72.2	3.2	312	10 BG230513	BG230513 na144a06-
34	72	3.1	416	9 AM835797	AM835797 OVA-LT001
35	72	3.1	764	10 B1915216	B1915216 603184745
36	71.8	3.1	2032	11 AF161549	AF161549 Homo sapi
37	71.4	3.1	794	10 BF983586	BF983586 602306575
38	71.2	3.1	327	9 A1344682	A1344682 qp09h04.x
39	71.2	3.1	363	9 A1358986	A1358986 qy19h02.x
40	71.2	3.1	372	9 A1370636	A1370636 ta38g12.x
41	71.2	3.1	380	9 A1244892	A1244892 qj98f05.x
42	71.2	3.1	380	9 A1298789	A1298789 qm91c02.x
43	71.2	3.1	388	9 A1291781	A1291781 qm75b05.x
44	71.2	3.1	393	9 A1214952	A1214952 qm31d12.x
45	71.2	3.1	394	9 A1439755	A1439755 t161a07.x

## ALIGNMENTS

RESULT 1  
LOCUS B1262402 480 bp mRNA linear EST 17-JUL-2001  
DEFINITION 602953836T1 NIH\_MGC\_99 Homo sapiens CDNA clone IMAGE:5088033 3',  
mRNA sequence.  
ACCESSION B1262402.1 GI:14822582  
VERSION B1262402.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.  
1 (Bases 1 to 480)  
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
AUTHORS Unpublished (1999)  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA Library  
Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM1846 row: 9 column: 10  
High quality sequence start: 27  
High quality sequence stop: 480.  
Location/Qualifiers  
1. 480  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5088033"  
/clone\_1ib="NIH\_MGC\_99"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph. Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GCCACGAG(C). Size-selected >500bp for average insert size



DB 149 CCAGAGCAATACATACAGGAGGAGCTAGCTATGTTCCAAATAGAACTTATTATTAATAA 208  
QY 1070 cggaaacctgaaataatgctc 1091  
||| ||| |||  
DB 209 AACAAATGCTCACTATAGTTT 230

RESULT 4  
A0695956 521 bp DNA linear GSS 06-JUL-1999  
LOCUS A0695956  
DEFINITION HS.2160.A2.F03.T7C.C1F Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate-2160 Col-6 Row-K, DNA sequence.  
ACCESSION A0695956  
VERSION A0695956.1 GI:5386204  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 521)  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
9380589  
JOURNAL Contact: Mahalir, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holman, T.,  
MEDLINE University of Washington  
COMMENT High Throughput Sequencing Center  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
BAC end Web Server: <http://www.htsc.washington.edu>  
Plate: 2160 Row: K column: 6  
Seq primer: 17  
Class: BAC ends  
High quality sequence stop: 521.

FEATURES  
Location/Qualifiers  
1..521  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate-2160 Col-6 Row-K"  
/clone\_id="C1F Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in  
E-Coli DH10B"

BASE COUNT 150 a 128 c 92 g 148 t 3 others  
ORIGIN

Query Match 3.4%; Score 78.4; DB 12; Length 521;  
Best Local Similarity 73.5%; Pred. No. 3.3e-09;  
Matches 100; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 906 gcaaacctctcgttaaaacacagagaaactttagctctgtgcatatgct 965  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 290 GCAAACTTTTCTGTAAGGCGCAGATAGTAATATTGTACACTTTCGAGACATATGCT 349  
QY 966 ctgaagcaactactcactctgctctgtgtagcagaagcaattgcaacaatagtc 1025  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 350 CTCCTGTTCCAACTTACTCTGCGTTGTAGCACAAAGCAGCCATAGCAATATGCA 409  
QY 1026 aacaacatatgtgac 1041  
||||| |||||||  
DB 410 GACAAATGCATGTGAC 425

RESULT 5  
AM183569 435 bp mRNA linear EST 18-NOV-1999  
LOCUS

DEFINITION xj76h04.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2663191 3', mRNA sequence.  
ACCESSION AM183569  
VERSION AM183569.1 GI:6452083  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 435)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -400P from Glibco  
High quality sequence stop: 420.

FEATURES  
Location/Qualifiers  
1..435  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="IMAGE:2663191"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site.1: Not 1; Site.2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBL19W, testis NHT, and B-cell  
NCL-GAP-GCBI) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 149 a 82 c 60 g 144 t  
ORIGIN

Query Match 3.4%; Score 78.2; DB 9; Length 435;  
Best Local Similarity 59.6%; Pred. No. 3.6e-09;  
Matches 201; Conservative 0; Mismatches 108; Indels 28; Gaps 3;

QY 901 gttggcaactcttctcgttaaaacacagagaaactatttagctctgtgcatatg 960  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 108 GGTGAGCAAACTTTTCTGTAAGGCGCAGATAGTAATATTGTACACTTTCGAGACAT 167  
QY 961 atggtctcagtcacactactactctctgctctgtgtagcagaagcaattagcaaat 1020  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 168 TTGGCTTTTGTGAACTTACTCAACTCTGCTAACATMACACAAAAGCTGCCAGACAACT 227  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
QY 1021 atgtcaacaacacatagtgaccccatgaaacttattat---taggataaggaacc 1077  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 228 ATGTAAACAATGAGGCTGACTGTAATAAATGTAATTTATGACACGAAATTTGAACT 287  
QY 1078 tgaataataagctcttctc-----tttgattttccccaatcaat 1121  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 288 TCAAAATCATTTTCATGCTTCCCAAAACATTTGACTTTTGACATTTTTCACACATTAA 347  
QY 1122 aacgtaaactactcttagtcgcgaaggttaagccatctcagctcagtgagcagc 1181  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 348 AAAGTAAACCAATCTTAGGTGACAGGCTGTGAAACAACT-----AGTCATAGGC 398  
QY 1182 tggattggcttggacctagctgagccatccctg 1218  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
DB 399 TTAATCTGGCCACAGCCATAGTTGCCAATGCCG 435

RESULT 6



LOCUS	A1128823	393 bp	mRNA	linear	EST 11-SEP-1998
DEFINITION	q94d06.s1 Soares_fetal_heart_NbH19w Homo sapiens cDNA clone				
ACCESSION	IMAGE:1694411 3', mRNA sequence.				
VERSION	A1128823				
SOURCE	A1128823.1	GI:3597337			
ORGANISM	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	N1 (bases 1 to 393)				
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/cgap.				
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
FEATURES	Unpublished (1997)				
source	Contact: Robert Strausberg, Ph.D. Email: cgaps-femail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Seq primer: -40ml3 fwd. ET from Amersham.				
	Location/Qualifiers				
	1..393				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_id="IMAGE:1694411"				
	/clone_lib="Soares_fetal_heart_NbH19w"				
	/sex="unknown"				
	/dev_stage="19 weeks"				
	/lab_host="DH10B (ampicillin resistant)"				
	/note="Organ heart; Vector: pF773D (Pharmacia) with a modified polylinker; Site_1: Not I - Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGAAGTGAGGGGCGGCCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pF773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Patino Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbH19w."				
BASE COUNT	140 a	70 c	53 g	130 t	
ORIGIN					
Query Match	3.4%	Score 77.4;	DB 9;	Length 393;	
Best Local Similarity	60.2%;	Pred. No. 5.6e-09;			
Matches 174;	Conservative 0;	Mismatches 96;	Indels 19;	Gaps 2;	
Oy	901 gttagggaacactctccgcgttaagaaagcaagaacgaattttagggctcgttgcat	960			
Dd	103 GGTTCGCAAACTTTTCTGTATAAGGGCCATTAACAATATTTCAGTCTGTGACAGCAT	162			
Oy	961 atggtctcaagtacaactactcatctctgcgcctctgtagcaagaaagcaattagaacaat	1020			
Dd	163 TTGGTCTTTGGTGAAGACTCCTCAACTGTGCTAACCTAATACAAAAGCTGCCACAGCAAT	222			
Oy	1021 atgtcaacaacatatgtgacccccatgaaaacttaattat---tatgataaggaaac	1077			
Dd	223 ATGTAAACAAATGAGGGTGACTGTATATAAATGTTATTTATGACACTGAAATTTGAAC	282			
Oy	1078 tgaaaaataatgctctctc-----ttgatltttcccacatcataaa	1121			
Dd	283 TCMAATCATTTTCATGCTCCCAAAAACATTTCACCTTTTGACATATTTTCAACCATTTAA	342			
Oy	1122 aaacgtaaaaactactcttagtcgaagtgtaagccaatctcgaacttg	1170			
Dd	343 AAAGTAAAAACCATTTCTTAGGTGACAGGCTGTAGAAAAATAGTCAATAG	391			
RESULT	9				
LOCUS	W76494/c	411 bp	mRNA	linear	EST 17-OCT-1996

[illegible]

Qy	1122	aacgtttaaactactcttagtcgcgaaggtttagccatcttaagcttag	1170
Db	85	AAAGTAAAAACCATTTCTTAGTGACACGCTGTGAAAAAACTAGTCATAG	37
RESULT	10		
AG154552			
LOCUS			
DEFINITION			
ACCESSION	AG154552	680 bp	DNA
VERSION	AG154552.1		Linear
KEYWORDS	GSS: GSS (genome survey sequence).		GSS 09-JAN-2002
SOURCE	Pan troglodytes male lymphocytes DNA, clone_1lb:RPCL-43 Chimpanzee Male BAC library clone:RP43-018J16.TJ.		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
AUTHORS	1 (sites)		
TITLE	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,		
JOURNAL	Totoki,Y., Watanabe,H. and Sakaki,Y.		
REFERENCE	BAC end sequences of Library RPCL-43		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 680)		
JOURNAL	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,		
REFERENCE	Totoki,Y., Watanabe,H. and Sakaki,Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical		
JOURNAL	and Chemical Research (RIKEN), Genomic Sciences Center (GSC);		
REFERENCE	1-7-22 Shuhiro-Chou,Tsurumi-Tsu, Yokohama, Kanagawa 230-0045, Japan		
AUTHORS	(E-mail:climbsegasc-riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,		
TITLE	Tel:81-45-503-9111 Fax:81-45-503-9170)		
JOURNAL	Clones are derived from the chimpanzee BAC library RPCL-43 This BAC		
REFERENCE	end was generated during the R&D process and may have higher chance		
AUTHORS	of clone tracking errors.		
TITLE	PRIMERS		
JOURNAL	Sequencing: TJ		
REFERENCE	LIBRARY		
AUTHORS	Vector : pBACe3.6		
TITLE	R.Site 1 : ECORI		
JOURNAL	R.Site 2 : ECGRI		
REFERENCE	Location/Qualifiers		
AUTHORS	1..680		
TITLE	/organism="Pan troglodytes"		
JOURNAL	/db_xref="taxon:9598"		
REFERENCE	/clone="RP43-018J16.TJ"		
AUTHORS	/sex="male"		
TITLE	/cell_type="lymphocytes"		
JOURNAL	/clone_1lb="RPCL-43 Chimpanzee Male BAC Library"		
REFERENCE	207 a 127 c 133 g 213 t		
AUTHORS	BASE COUNT		
TITLE	ORIGIN		

[illegible]

RESULT	11
AI92674	
LOCUS	AI92674 400 bp mRNA linear EST 02-SEP-1998
DEFINITION	w03g06.x1 NCI-CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2458138 3'
ACCESSION	AI92674
VERSION	AI92674.1 GI:5662162
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 400)
JOURNAL	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://www.bio.lnl.gov/bdrp/image/image.html">www.bio.lnl.gov/bdrp/image/image.html</a> Seq primer: -40UP from GIDCO High quality sequence stop: 382. Location/Qualifiers 1..400 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="IMAGE:2458138" /clone_lib="NCI CGAP Gas4" /tissue_type="poorly differentiated adenocarcinoma with signed ring cell features" /lab_host="DH10B" /note="Organ: stomach; Vector: pCMV-SPOUT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"
BASE COUNT	133 a 73 c 54 g 140 t
ORIGIN	
Query Match	3.4%; Score 77.2; DB 9; Length 400;
Best Local Similarity	64.9%; Pred. No. 6.3e-09;
Matches 131; Conservative	0; Mismatches 68; Indels 3; Gaps 1;

LOCUS	DEFINITION	AI240516	322 bp	mRNA	linear	EST 30-NOV-1998
RESULT 12		AI240516/c				
Db	127	GGTCAGCAAACTTTTCTGTGTAAAGGGCCATACAAATACATATTTCAGTCTCTGCAGACAT	186			
Qy	961	atggtctcagtcacaactcattcctctgctcctcgtagcagcaagaacatlagcaaat	1020			
Db	187	TTGGTCTTTTGCTGAACCTACTCACTCGCTGACATTAACATAACAAAAAGCTGCCACAGACAAT	246			
Qy	1021	atgtcacacaacatcgtcaccccatgaaacttattatctatgtgacagcaacactga	1080			
Db	247	ATGTAAACAAAGAGGGGTGACGTGTAAATTAATTTTAT---TTATGAACACTGAAATTTTGA	303			
Qy	1081	aaataatgctctctcttcttgatc	1102			
Db	304	ACTTCAAAATCATTTTCATGCTT	325			

TITLE	Kucaba,T., Lacey,M., Le,N.,Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.			
JOURNAL	WashU-Werck EST Project 1997			
COMMENT	Unpublished (1997)			
	Contact: Wilson RK			
	Washington University School of Medicine			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: estevatson.wustl.edu			
	This clone is available royalty-free through LBNL; contact the			
	IMABO Consortium (info@image.llnl.gov) for further information.			
	Seq primer: -41m13 fwd. ET from Amer sham			
	High quality sequence stop: 321.			
FEATURES	location/Qualifiers			
source	1..386			
	/organism="Homo sapiens"			
	/db_xref="GDB:5977767"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:757156"			
	/clone_lib="Soares_testis_NHT"			
	/sex="male"			
	/lab_host="DH10B"			
	/note="Vector: pT73D-Pac (Pharmacia) with a modified			
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA			
	was prepared from mRNA obtained from Clontech Laboratories			
	, Inc., and primed with a Not I - Oligo(dT) primer [5'			
	TGTATCCAACTGGAAGGAGGAGGCGGCCCATTTTTTTTTTTT 3'].			
	Double-stranded cDNA was ligated to Eco RI adaptors			
	(Pharmacia), digested with Not I and cloned into the Not I			
	and Eco RI sites of the modified pT73 vector. Library			
	went through one round of normalization to Cot5, and was			
	constructed by Bento Soares and M. Fatima Bonaldo. "			
BASE COUNT	111 a 72 c 68 g 135 t			
ORIGIN				
Query Match	3.4%; Score 77; DB 9; Length 386;			
Best Local Similarity	63.0%; Pred. No. 7e-09;			
Matches 119; Conservative	0; Mismatches 70; Indels 0; Gaps 0.			
QY	881	tttcccatgctgcgtgaccagcaggtgtggcgaaacatcttcctgtaaagaacagcaggaacta	940	
Db	139	TTTTTTCTATCTAGAACATGAGGCTTGCCAAACTTTTCTATATAAGGCGCTGTATATAATA	198	
QY	941	ttttagcctctgtgtgccatagtcctcagtcacacactcactcctcgtcctctgtagca	1000	
Db	199	TTTTCAGGCTTTGAGGGTTCACATGTCCTCTGTGATGACACTACTCAATTCGCGCATTAATGTA	258	
QY	1001	cgaagcgaattagcacaacatatgtcaacaacatatgtgaccccatgaaaaactttat	1060	
Db	259	CAGAACTGCGCAACAGCAATACGTCAAAACAATAAGAGTGTGTGCCAATAAACAATT	318	
QY	1061	ttatgatga 1069		
Db	319	ATTTCACAA 327		
RESULT 14				
AA444117/c	393 bp mRNA linear EST 03-JUN-1997			
LOCUS	zv51c03.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:757156			
DEFINITION	5', mRNA sequence.			
ACCESSION	AA444117			
VERSION	AA444117.1 GI:2156792			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,			
AUTHORS				





---

1  
2  
3

4  
5  
6

7



```
QY 1076 cctgaataaataatgctctctctctctctctctctctcccaatcatlaaaacgtaaaactac 1135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 TTTTAATTTTCATATTTATTTCTTTTGTGATATTTTGAACACATTTTAAATAAAGAAACA 418
QY 1136 tcttagtcgaaggttaagcaatctcagcttagcagctgagcagctgagctgagctgagct 1195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 TTCTTGCTGCACAGGCTATGCA-----TAACAGTCAGCAAGCCAGATTTGGCTCAT 366
QY 1196 gacctacagctggccaatcccgatt 1221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 AGGCCACAGTATGCCAACCCCTGTTT 340

RESULT 2
US-10-027-632-119946/c
; Sequence 119946, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 119946
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-119946

Query Match 3.6%; Score 82.8; DB 7; Length 1007;
Best Local Similarity 58.0%; Pred. No. 3,3e-12;
Matches 189; Conservative 0; Mismatches 127; Indels 10; Gaps 2;

QY 896 cccaggttgagcaactctcctgtaaaagacagagcaactatcttaagctctgctg 955
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 CCAAGGGGTAGTTAACTTTTCTGTAAGAGCAGACGATTAATATTTTAAAGCTTTGTG 596
QY 956 gccatatggtctcagtcacaactatcctcctctgtagcagcaagaagaactagca 1015
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GCCATAGGCTCTGTGCGATCTATCACTTTGCTGTGTCAGCAAAATGCAGCCATG 536
QY 1016 acaatatgccaacaaacatatgaccccaatgaaaaacttattattatgagtaagaa 1075
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 ATTAATACATAAATGAATGAGATGTGTTCAATTAACAT--TTATTTATGAGACACTGAAA 478
QY 1076 cctgaataaataatgctctctctctctctctctctcccaatcatlaaaacgtaaaactac 1135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 TTTTAATTTTCATATTTATTTCTTTTGTGATATTTTGAACACATTTTAAATAAAGAAACA 418
QY 1136 tcttagtcgaaggttaagcaatctcagcttagcagctgagcagctgagctgagctgagct 1195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 TTCTTGCTGCACAGGCTATGCA-----TAACAGTCAGCAAGCCAGATTTGGCTCAT 366
QY 1196 gacctacagctggccaatcccgatt 1221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 365 AGGCCACAGTATGCCAACCCCTGTTT 340

RESULT 3
US-10-105-299-12107/c
; Sequence 12107, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12107
; LENGTH: 37437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12107

Query Match 3.6%; Score 81.8; DB 7; Length 37437;
Best Local Similarity 62.9%; Pred. No. 3.2e-11;
Matches 144; Conservative 0; Mismatches 82; Indels 3; Gaps 1;

QY 807 actccctctctctctcattccatctcagctgagcaggtatcctgctcccaagctca 866
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32330 AGTCCACTTCATTTCTTATGCTACTTCTTCAGATATATTTCTTAATAAACAACATPA 32271
QY 867 ca---tctctactccgcttcccatgctgacccaagcttgagcaactctctctgtaa 923
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32270 AAATGTGCTTTGCAATGCTCAGATATCTAACAACAGGGGTGACGAAACTTTTCTGAAA 32211
QY 924 gaaccagacgaagaactatcttagctctgctgtagcagctatgctcagctcacaactactca 983
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32210 GAGCCAGATGTGAATATTTTAAAGCTTTGTGGCCAGCATCTCTGTTGCAACTATATCA 32151
QY 984 tctctgctctgtagcagcaagaactagcaacatatgccaacaac 1032
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32150 TTCTGCTGCTATAGTACAAAAGCAGCTATAAACAATTAATTAACAAAC 32102

RESULT 4
US-10-105-299-12108/c
; Sequence 12108, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12108
; LENGTH: 37442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12108

Query Match 3.6%; Score 81.4; DB 7; Length 37442;
Best Local Similarity 63.0%; Pred. No. 4.1e-11;
Matches 143; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

QY 809 tctcctctctctctcattccatctcagctgagcaggtatcctgctcccaagctcaca 868
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32329 TCCACTTCATTTCTTCACTGCTACTTCTTCCAGTATATATTTCTTAATAAACAACATPAAA 32270
QY 869 ---tctctactccgcttcccatgctgagcagcaagcttgagcaactctctgtaaga 925
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32269 ATTTGCTTTTGCATATGCTCAGATATCTAACAACAGGGGTGACAAACTTTTCTGAAAAGA 32210
```

```
OY      926   accagcaaggaaactatttttaggcctcgtgftgcacataigtcctaagtaaccaactcatc 965
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     32209 GCCAGATGGTAAATTATTTTAGGCTTGtggGCCAACAFTTCtCGGTGCACACTVATTCATT 32150

OY      986   tctgcctcttgacacgaagaacatlagacaacaatgatgcacaanaac 1032
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     32149 TCTGCTCCTATTAGTACAAAAGCATATAAACAAATAAGTAAGCAAAC 32103


RESULT       5
US-09-629-469A-12215/c
; Sequence 12215, Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGITAMA, TOMOTASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OHSUKI, TEISUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 08435/0123
CURRENT APPLICATION NUMBER: US/09/629,469A
PRIOR APPLICATION FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 12215
LENGTH: 5294
TYPE: DNA
ORGANISM: Homo sapiens
US-09-629-469A-12215


Query Match              3.5%; Score 81; DB 5; Length 5294;
Best Local Similarity    74.5%; Pred. No. 2.2e-11;
Matches 102; Conservative 0; Mismatches 35; Indels 0; Gaps 0

OY      894   gaccacagtggccaacctccctcgftraagaagcacagcagaaccttttaggccttgt 953
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     1915   GACCAGAGAAACAAAAAATCTTTTCGTAAAGGCGCACGAAGAATATTATTTAGGTTTGT 1856

OY      954   gtgcacatatgtctcagtcacacactactcatctcgtcctctgttagcacgaagaatlag 1013
          |  |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     1855   GGCGCAATATGGTCTCTGTGGCAATTACTTAACCTCTGCCATCATATAGCAGAAGAAGCAGCTAC 1796

OY      1014  caacaataatgtcaaca 1030
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     1795   AGACCAATATGTACGCAA 1779


RESULT       6
US-09-918-995-12893/c
; Sequence 12893, Application US/09918995
GENERAL INFORMATION:
```

```

1  APPLICANT Hyseq, Inc.
2  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
3  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
4  FILE REFERENCE: 20411-756
5  CURRENT APPLICATION NUMBER: US/09/918,995
6  CURRENT FILING DATE: 2001-07-30
7  PRIOR APPLICATION NUMBER: US/09/735,076
8  PRIOR FILING DATE: 1999-01-20
9  NUMBER OF SEQ. ID NOS: 38054
10 SOFTWARE: FASTSEQ for Windows Version 3.0
11 SEQ ID NO 12893
12 LENGTH: 503
13 TYPE: DNA
14 ORGANISM: Homo sapiens
15 FEATURE:
16 NAME/KEY: misc_feature
17 LOCATION: (1)..(503)
18 OTHER INFORMATION: n = A,T,C or G
19 US-09-918-995-12893
20
21 Query Match 3.5%; Score 79.4; DB 5; Length 503;
22 Best Local Similarity 58.4%; Pred.No.2,1e-11;
23 Matches 157; Conservative 0; Mismatches 111; Indels 1; Gaps 1
24
25 Qy 887 atgcctgagcccgaggttggcgaacctcttcctgtaagaagacagacgaagacttattag 946
26 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 Db 459 ATGCGTTCCAAATCCATCGGTGGTCACGAACTTGTATATGCGCCACAGTAAGTAATATTAG 400
28
29 Qy 947 gctcttgtyccatataatgctcagtcacaactcaattctctgtcctctgtagcagaag 1006
30 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 Db 399 GCTCTGTGGGCGCACACATCTTGTGTACAGCTACTCAGCTGTGCCCTGTACACGACAG 340
32
33 Qy 1007 caattgcaacaatattgcaacaacataatgtaaccacgaagaaacttattattatcg 1066
34 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 Db 339 CAGTCACAGGATATATGTGATGAATGAATGATGATGTCGCCCATTTCTTAACATTTTCATA 280
36
37 Qy 1067 atacggaacctg-aaataatgctcttccttcttgaatttttcccaatcatlaaaaaag 1125
38 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 Db 279 ACATATGACATGTCACAAAATATATTTCTTTGGATTTTTTTTCAGTCATTTAAAGSTA 220
40
41 Qy 1126 taataactactcttagtgcgaagttaa 1154
42 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 Db 219 TAAATAACACAAACATGAAAAAATCTTAA 191
44
45 RESULT 7
46 US-09-785-276A-55453/C
47 Sequence 55453, Application US/09785276A
48 GENERAL INFORMATION:
49 APPLICANT: Schlegel, Robert
50 APPLICANT: Endege, Wilson
51 APPLICANT: Monahan, John
52 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
53 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
54 TITLE OF INVENTION: HUMAN PROSTATE CANCER
55 FILE REFERENCE: MFI-007B
56 CURRENT APPLICATION NUMBER: US/09/785,276A
57 CURRENT FILING DATE: 2001-02-16
58 PRIOR APPLICATION NUMBER: 60/183,319
59 PRIOR FILING DATE: 2000-02-17
60 PRIOR APPLICATION NUMBER: 60/189,862
61 PRIOR FILING DATE: 2000-03-16
62 PRIOR APPLICATION NUMBER: 60/207,454
63 PRIOR FILING DATE: 2000-05-25
64 PRIOR APPLICATION NUMBER: 60/211,314
65 PRIOR FILING DATE: 2000-06-09
66 PRIOR APPLICATION NUMBER: 60/219,007
67 PRIOR FILING DATE: 2000-07-18
68 PRIOR APPLICATION NUMBER: 60/255,281
69 PRIOR FILING DATE: 2000-12-13
70 NUMBER OF SEQ ID NOS: 62232
71 SOFTWARE: FastSEQ for Windows Version 4.0

```

```

; SEQ ID NO 55453
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 561..582
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-55453
```

```

Query Match          3.5%: Score 79.4; DB 5; Length 627;
Best Local Similarity 67.7%: Pred. No. 2.3e-11;
Matches 128; Conservative 0; Mismatches 56; Indels 5; Gaps 1;
```

```

Qy 907 caaaccttcctgtaagaacagacagaaactatttgctgctgctgacatgtgc 966
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 476 CGAACTTTCTGTAAAGAGCCAGAGTAATTTTCAGACTTTGTAGGCCAACACTC 417
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 967 tcaatcacactacatcatctgctctgtagacagaagaatagacacatatgtca 1026
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 TCTGTCAAAATATTCATCTGCGCATTTGTCCAGAAAGCAGCCAGCAATACACA 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1027 acaaacatatgtgac-----ccatgaaacatttatcttctgtagatcgagaaacctgaa 1081
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 ACAAGTAGACCTGCGCTGTCCAAATAAACTTATTTCAAACTAGAGCTAAGCGTGAT 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1082 aataatgtc 1090
      |||||
Db 296 GATCGAGTC 288
      |||||
```

```

RESULT 8
US-10-027-632-151080
; Sequence 151080, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151080
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151080
```

```

Query Match          3.5%: Score 79; DB 7; Length 749;
Best Local Similarity 72.0%: Pred. No. 3.2e-11;
Matches 103; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

```

Qy 885 ccatgctgagccagcaggttgaggcaactctctcttaagaagccagacgaactatttt 944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 cctggtttagccaggttgaggcaactctctcttaagaagccagacgaactatttt 431
```

```

Qy 945 aggcctgtgcatatgctcagtcacaaactactatctctgctctgtagacgaa 1004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 agactctgtggccatttgccctctgctcagacaaactactatctctgctgtagacgaa 491
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1005 agcaatagcaacatatgtcaa 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 accagccacagacatatgtcaa 514
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 9
US-10-027-632-151081
; Sequence 151081, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151081
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151081
```

```

Query Match          3.5%: Score 79; DB 7; Length 749;
Best Local Similarity 72.0%: Pred. No. 3.2e-11;
Matches 103; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

```

Qy 885 ccatgctgagccagcaggttgaggcaactctctctgtaagaagccagacgaactatttt 944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 cctggtttagccaggttgaggcaactctctctgtaagaagccagacgaactatttt 431
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 945 aggcctgtgcatatgctcagtcacaaactactatctctgctctgtagacgaa 1004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 agactctgtggccatttgccctctgctcagacaaactactatctctgctgtagacgaa 491
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1005 agcaatagcaacatatgtcaa 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 accagccacagacatatgtcaa 514
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 10
US-10-027-632-7748/C
; Sequence 7748, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
```

```

RESULT 13
US-10-027-632-142654/c
; Sequence 142654, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 10887, 129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

```

```

1 PRIOR FILING DATE: 2000-03-29
2 PRIOR APPLICATION NUMBER: US 60/185,218
3 PRIOR FILING DATE: 2000-02-24
4 PRIOR APPLICATION NUMBER: US 60/167,363
5 PRIOR FILING DATE: 1999-11-23
6 PRIOR APPLICATION NUMBER: US 60/156,358
7 PRIOR FILING DATE: 1999-09-28
8 PRIOR APPLICATION NUMBER: US 60/146,002
9 PRIOR FILING DATE: 1999-08-09
10 NUMBER OF SEQ ID NOS: 325720
11 SOFTWARE: FastSeq for Windows Version 4.0
12 SEQ ID NO 142654
13 LENGTH: 538
14 TYPE: DNA
15 ORGANISM: Human
16 US-10-027-632-142654

```

Query Match	3.4%	Score	76.6	DB 7	Length	538
Best Local Similarity	64.7%	Pred. No.	1.3e-10			
Matches 130; Conservative	0	Mismatches	69	Indels	2	Gaps 1

QY	901	gttgaggaaactcttctctgtaaaagacagaaagaaacattttagtgctcgtgtagcat	960
DB	214	GTTTGGCAAACTCTTCTCTTAAAGGACCAAGTAGTAACATTTTTAGGTTTTGCAAGCCAC	155
QY	961	atggtctcagtcacaaactactacatctctgctctgtgacgaagaaacattagcaacat	1020
DB	154	ATGTCCTCTGTACGAACTACTACTACCTCTGCGACCTGTAGCTCTAAAGCAGCCACAGACAT	95
QY	1021	atgtcaacaacacatctgtacccacagaaacttattt--attatgatacggaaacat	1078
DB	94	ATTTAATTAATGACAGAGTGCGCGCTCTTCATTAAAACTTCATTGTGTACACTGAATTT	35
QY	1079	gaaataatgtcttctcttcttgg	1099
DB	34	GAAATTATATTAATTTTCATG	14

RESULT 14  
US-09-785-276A-55307/c  
; Sequence 55307, Application US/09785276A

? APPLICANT: Schlegel, Robert  
 ? APPLICANT: Endegat, Wilson  
 ? APPLICANT: Monahan, John  
 ? TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ? TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
 ? TITLE OF INVENTION: HUMAN PROSTATE CANCER  
 ? FILE REFERENCE: MRI-007B

```

1 CURRENT APPLICATION NUMBER: US/09/785,276A
2 CURRENT FILING DATE: 2001-02-16
3 PRIOR APPLICATION NUMBER: 60/183,319
4 PRIOR FILING DATE: 2000-02-17
5 PRIOR APPLICATION NUMBER: 60/189,862
6 PRIOR FILING DATE: 2000-03-16
7 PRIOR APPLICATION NUMBER: 60/207,454
8 PRIOR FILING DATE: 2000-05-25
9 PRIOR APPLICATION NUMBER: 60/211,314
10 PRIOR FILING DATE: 2000-06-09
11 PRIOR APPLICATION NUMBER: 60/219,007
12 PRIOR FILING DATE: 2000-07-18
13 PRIOR APPLICATION NUMBER: 60/255,281
14 PRIOR FILING DATE: 2000-12-13
15 NUMBER OF SEQ ID NOS: 62232
16 SOFTWARE: FastSeq for Windows Version 4.0
17 SEQ ID NO 55307
18 LENGTH: 608
19 TYPE: DNA
20 ORGANISM: Homo sapiens
21 US-09-785-276A-55307

```

```

Query Match      3.4%; Score 76.6; DB 5; Length 608;
Best Local Similarity 74.0%; Pred. No. 1.3e-10;
Matches 97; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```

Oy		961	atgtgcacgtagcaactactctcctccgccctgtagaaagaagcatatgacaacat	1020
Dd		123	AAGTCCTTTTGCACATTAATCATCTCTCCCAATTTTTGTGACAAGCAGCCATTAGACAAAT	64
Oy		1021	atgtccaacaaa	1031
Dd		63	GTAATAAACAAA	53

```

RESULT 15
US-10-027-632-233742/C
; Sequence 233742, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,356
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233742
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-233742

```

Query Match	3.3%	Score 75.6;	DB 7;	Length 556;
Best Local Similarity	65.9%;	Pred. No. 2.4e-10;		
Matches 108; Conservative	1;	Mismatches 55;	Indels 0;	Gaps 0;

Qy	901	gttgggaacactcttcctgtgaagaacacagacggagacatttttgccctctgtgtgcac	960
Db	472	GGTGGCAAAATTTTCTGTGAAAAACCTTGATGTAATAATTTTARGCTTGCACAGTAT	413
Qy	961	atgtttccagtcgaacaattactcaactcttcctctctgttagcagaagaagcaattgacacac	1020
Db	412	ATGTTCTCTATCCACACTACTGTGCTCAGCCACTGTACGAGAAAACGATTTGTAAACAT	353
Qy	1021	atgtcaacaacaacatattgtgagcccaactgtaaacttattattat	1064
Db	352	ATGTAACTGAAATGAAACATGGCTGTGTTCCAAATTAATTAATTT	309

Search completed: August 16, 2002, 05:55:47  
Job time: 11641 sec









FT	protein_bind	1109..1113	/tag= f	"C/EBP"
FT		/bound_moiety=		
FT	protein_bind	1207..1214	/tag= g	"AP1"
FT		/bound_moiety=		
FT	protein_bind	1209..1214	/tag= h	"C/EBP"
FT		/bound_moiety=		
FT	protein_bind	1530..1537	/tag= i	"AP1"
FT		/bound_moiety=		
FT	protein_bind	2115..2121	/tag= j	"AP1"
FT		/bound_moiety=		
FT	protein_bind	2126..2133	/tag= k	"C/EBP"
FT		/bound_moiety=		
FT	TATA_signal	2210..2214	/tag= l	
XX				
PN	MO200164921-A1.			
XX				
XX	07-SEP-2001.			
XX				
PE	28-FEB-2001; 2001MO-US06782.			
XX				
PR	29-FEB-2000; 2000US-0515369.			
PA	(UYCO ) UNIV COLUMBIA NEW YORK.			
P1	Fisher PB, Madireddi MT:			
XX				
XX	WPI: 2001-565508/63.			
PT				
PT	Melanoma differentiation associated gene-7 promoter capable of			
PT	treating cancer comprises directing transcription of heterologous			
PT	coding sequence encoding tumour suppressor polypeptide positioned			
PT	downstream, useful for treating cancer			
PS	Claim 1; Fig 6A-B; 132pp; English.			
XX				
CC	The present sequence is that of the promoter region of the human			
CC	melanoma differentiation associated gene-7 (Mda-7). It includes			
CC	2240 bp from the 5' flanking region of the Mda-7 gene. The			
CC	promoter was isolated from a human placental genomic library			
CC	using a PCR-based method. The Mda-7 promoter exhibits melanocyte			
CC	tissue specificity and can only be activated in the targeted			
CC	tissue, i.e. the skin. Therefore, a gene of interest driven by			
CC	the Mda-7 promoter will be differentially expressed in these cells,			
CC	mitigating systemic toxicity. A recombinant expression construct			
CC	in which the human Mda-7 promoter is operably linked to a coding			
CC	sequence encoding a tumour suppressor protein is claimed. The			
CC	tumour suppressor is preferably p21, retinoblastoma protein or p53.			
CC	A host cell comprising the expression construct is also claimed,			
CC	and is preferably a tumour cell selected from a melanoma,			
CC	neuroblastoma, astrocytoma, glioblastoma multiforme, cervical			
CC	cancer, breast cancer, lung cancer or prostate cancer cell.			
CC	The promoter preferably comprises nucleotides 1-2240 of the present			
CC	sequence. A pharmaceutical composition including the recombinant			
CC	expression construct is used in a claimed method of treating			
CC	melanoma, neuroblastoma, astrocytoma, glioblastoma multiforme,			
CC	cervical cancer, breast cancer, colon cancer, prostate cancer,			
CC	osteosarcoma, chondrosarcoma or a cancer of the central nervous			
CC	system.			
XX				
Sequence	2286 BP; 547 A; 592 C; 510 G; 637 T; 0 other;			

Query Match	100.0%;	Score 2286;	DB 22;	Length 2286;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2286;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	taataagactactaataagggctgcagctcogatacactctttgaacccaggtctgcgctc	60
Db	1	taataagactactaataagggctgcagctcogatacactctttgaacccaggtctgcgctc	60
QY	61	ccaagaactgttactcataactatagatctccaactgtgtgttgccaaagttccactagttct	120
Db	61	ccaagaactgttactcataactatagatctccaactgtgtgttgccaaagttccactagttct	120
QY	121	ctctctgaccttcctctcgtgaagtaabaatgtctaagataagctcaatcgggaagctgaagccc	180
Db	121	ctctctgaccttcctctcgtgaagtaabaatgtctaagataagctcaatcgggaagctgaagccc	180
QY	181	aggacacattgtttgctggaactatccatgftatataagatcccttcctccagacgaagtgagct	240
Db	181	aggacacattgtttgctggaactatccatgftatataagatcccttcctccagacgaagtgagct	240
QY	241	actcacagatccacaggtgtlaaccctgaagccagacgaagtgftatccatgaactcaatgtcctc	300
Db	241	actcacagatccacaggtgtlaaccctgaagccagacgaagtgftatccatgaactcaatgtcctc	300
QY	301	gttccagactgcgccccttaacagctcaatcccaactgcgctgcctcccgccctatcttgaga	360
Db	301	gttccagactgcgccccttaacagctcaatcccaactgcgctgcctcccgccctatcttgaga	360
QY	361	cagtagtctgaagatttcaagctgcgcggggggtcatttccctccaaactccctgcgtta	420
Db	361	cagtagtctgaagatttcaagctgcgcggggggtcatttccctccaaactccctgcgtta	420
QY	421	gctgtctcctgcgtcccaactcaactatctacccagacacttcaactgtgtctctcttcctg	480
Db	421	gctgtctcctgcgtcccaactcaactatctacccagacacttcaactgtgtctctcttcctg	480
QY	481	tctatacactgtccctctgacatactttatctacatagatgattgttaggtgtctctgtgaat	540
Db	481	tctatacactgtccctctgacatactttatctacatagatgattgttaggtgtctctgtgaat	540
QY	541	gcccttaaatccacaatgtgttggaaggggtgaggtgtgttggaagagatgtgcgtgttggtgt	600
Db	541	gcccttaaatccacaatgtgttggaaggggtgaggtgtgttggaagagatgtgcgtgttggtgt	600
QY	601	gctcaactctgtgaaggtgaagaactccggccctccagagaacaaagagattcaagctgtgtga	660
Db	601	gctcaactctgtgaaggtgaagaactccggccctccagagaacaaagagattcaagctgtgtga	660
QY	661	gctatagccaagacgagactgtctgcgcagggtgatttgcnaaagagattttgttgcctaaagaa	720
Db	661	gctatagccaagacgagactgtctgcgcagggtgatttgcnaaagagattttgttgcctaaagaa	720
QY	721	aatataacaacctgaatataagatgtgaagggaggggtgtgtgtgcgaagagatgtggaa	780
Db	721	aatataacaacctgaatataagatgtgaagggaggggtgtgtgtgtgcgaagagatgtggaa	780
QY	781	gagctgtgccaaggggtgtgttctactcaactccctcttctcttcaatctccactgaagctgtg	840
Db	781	gagctgtgccaaggggtgtgttctactcaactccctcttctcttcaatctccactgaagctgtg	840
QY	841	aggagagttatccctgtccccacgttcaactcttccactcccggttcccaatgtcgtggcccag	900
Db	841	aggagagttatccctgtccccacgttcaactcttccactcccggttcccaatgtcgtggcccag	900
QY	901	gtctgggcaaaactctctccctgttaagaagaacagaagagaaactattttaggtctgtgtgcatt	960
Db	901	gtctgggcaaaactctctctctgttaagaagaacagaagagaaactattttaggtctgtgtgcatt	960
QY	961	atggtctcagttacaaactactcatctctgtcctctgttagcagaagaagcaattagcaacaat	1020
Db	961	atggtctcagttacaaactactcatctctgtcctctgttagcagaagaagcaattagcaacaat	1020
QY	1021	atgtcaacaacaatatgtgaagcccaatgaanaacttttatataitgatacggaaacctga	1080
Db	1021	atgtcaacaacaatatgtgaagcccaatgaanaacttttatataitgatacggaaacctga	1080
QY	1081	aaataatgctcttctcttggatttttctcccaatcattataaanaagttaanaactactctta	1140

|||||  
Db 1081 aaataagctcttcttctgatttttccccaatcaataaagtaaaactactctta 1140  
Qy 1141 ggtcgaaggttaagccatctcagcttagcaatgtagcagctgattgtgctgtgacct 1200  
Db 1141 ggtcgaaggttaagccatctcagcttagcaatgtagcagctgattgtgctgtgacct 1200  
Qy 1201 acaagttagccaatccctgattcccaaaatglatctctcaaggatgtaggcaaatctat 1260  
Db 1201 acaagttagccaatccctgattcccaaaatglatctctcaaggatgtaggcaaatctat 1260  
Qy 1261 ggggaagtgctgatttaacaagagttaagaagcatcagaacttccagagcgttagcac 1320  
Db 1261 ggggaagtgctgatttaacaagagttaagaagcatcagaacttccagagcgttagcac 1320  
Qy 1321 atgcgaaggctctctaactacatccatctgattcaatctgtttcatggaagattctgcaag 1380  
Db 1321 atgcgaaggctctctaactacatccatctgattcaatctgtttcatggaagattctgcaag 1380  
Qy 1381 acaagaattccccaacctaagagctgtaggactgtgcttgtaggaacaactgtctgtctg 1440  
Db 1381 acaagaattccccaacctaagagctgtaggactgtgcttgtaggaacaactgtctgtctg 1440  
Qy 1441 atgcccccatctgggacatgtagttagaatctagaactgtagtgccttgctagctggaagatag 1500  
Db 1441 atgcccccatctgggacatgtagttagaatctagaactgtagtgccttgctagctggaagatag 1500  
Qy 1501 gtccagagctctgacgtcccttgccagtgctgagacatcaacgtctgtgttccctgagat 1560  
Db 1501 gtccagagctctgacgtcccttgccagtgctgagacatcaacgtctgtgttccctgagat 1560  
Qy 1561 ggttcagagccacaacagagccacaagactagccacaagagccacagcctccagcttctg 1620  
Db 1561 ggttcagagccacaacagagccacaagactagccacaagagccacagcctccagcttctg 1620  
Qy 1621 ggctgtgcatagtgtatcttcttcttcttctgtgttccaaactaaggagagggcac 1680  
Db 1621 ggctgtgcatagtgtatcttcttcttcttctgtgttccaaactaaggagagggcac 1680  
Qy 1681 atttggctgagtgatataaaccttagaggacatgtaggtagctgcatgtagcagaacaactc 1740  
Db 1681 atttggctgagtgatataaaccttagaggacatgtaggtagctgcatgtagcagaacaactc 1740  
Qy 1741 ctcaactcctgtgcctgtatgtagttaaaggagtagtactaagtgattattcttcgtgt 1800  
Db 1741 ctcaactcctgtgcctgtatgtagttaaaggagtagtactaagtgattattcttcgtgt 1800  
Qy 1801 ggaactatgtccagcagatagtagggcatcatattatgttaattttgacggagagagagac 1860  
Db 1801 ggaactatgtccagcagatagtagggcatcatattatgttaattttgacggagagagagac 1860  
Qy 1861 tggctgtagctgcacagtaataatgtaattatcatgtagcagaggtttacaagaacactct 1920  
Db 1861 tggctgtagctgcacagtaataatgtaattatcatgtagcagaggtttacaagaacactct 1920  
Qy 1921 gtgtgttttgccttgtttattacaacttgggacaatttttaaaattatatactgtag 1980  
Db 1921 gtgtgttttgccttgtttattacaacttgggacaatttttaaaattatatactgtag 1980  
Qy 1981 agagctgtagcagagagaagtagtagaacttggccccgccccacaacagccagtgtagagcc 2040  
Db 1981 agagctgtagcagagagaagtagtagaacttggccccgccccacaacagccagtgtagagcc 2040  
Qy 2041 tgaactcaaaacccaggtctcatctcaactcagagggctgtcttcccatgctgtatgttc 2100  
Db 2041 tgaactcaaaacccaggtctcatctcaactcagagggctgtcttcccatgctgtatgttc 2100  
Qy 2101 cttaaaagttagtggtgtagtaggaatgaaagtaattctctagaagaagcatgaccaatttcc 2160  
Db 2101 cttaaaagttagtggtgtagtaggaatgaaagtaattctctagaagaagcatgaccaatttcc 2160  
Qy 2161 ctcttcaccctcctcttcttctcaccctcccccatcaagccccatataatagccc 2220  
|||||

Db 2161 ctcttcaccctcctcttcttctcaccctcccccatcaagccccatataatagccc 2220  
Qy 2221 aaatctccacaagaagccttcttgctgctgcaaaccttactctgaatgacttccagcgt 2280  
Db 2221 aaatctccacaagaagccttcttgctgctgcaaaccttactctgaatgacttccagcgt 2280  
Qy 2281 gggagcg 2286  
Db 2281 gggagcg 2286  
RESULT 2  
AAK84165/c  
ID AAK84165 standard; DNA; 37437 BP.  
XX  
AC AAK84165;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38977.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.

PR	02-SEP-2000	2000US-0229513	
PR	06-SEP-2000	2000US-0230437	
PR	06-SEP-2000	2000US-0230438	
PR	08-SEP-2000	2000US-0231242	
PR	08-SEP-2000	2000US-0231243	
PR	08-SEP-2000	2000US-0231244	
PR	08-SEP-2000	2000US-0231414	
PR	08-SEP-2000	2000US-0231801	
PR	08-SEP-2000	2000US-0232080	
PR	12-SEP-2000	2000US-0231968	
PR	14-SEP-2000	2000US-0232927	
PR	14-SEP-2000	2000US-0232996	
PR	14-SEP-2000	2000US-0232999	
PR	14-SEP-2000	2000US-0232400	
PR	14-SEP-2000	2000US-0232401	
PR	14-SEP-2000	2000US-0233063	
PR	14-SEP-2000	2000US-0233064	
PR	14-SEP-2000	2000US-0233065	
PR	21-SEP-2000	2000US-0234423	
PR	21-SEP-2000	2000US-0234474	
PR	25-SEP-2000	2000US-0234697	
PR	25-SEP-2000	2000US-0234698	
PR	26-SEP-2000	2000US-0234699	
PR	27-SEP-2000	2000US-0235834	
PR	27-SEP-2000	2000US-0235836	
PR	02-OCT-2000	2000US-0237038	
PR	02-OCT-2000	2000US-0237039	
PR	02-OCT-2000	2000US-0237040	
PR	13-OCT-2000	2000US-0239935	
PR	13-OCT-2000	2000US-0239937	
PR	20-OCT-2000	2000US-0241260	
PR	20-OCT-2000	2000US-0241261	
PR	20-OCT-2000	2000US-0241185	
PR	20-OCT-2000	2000US-0241186	
PR	20-OCT-2000	2000US-0241807	
PR	20-OCT-2000	2000US-0241808	
PR	20-OCT-2000	2000US-0241809	
PR	01-NOV-2000	2000US-0244616	
PR	08-NOV-2000	2000US-0246474	
PR	08-NOV-2000	2000US-0246475	
PR	08-NOV-2000	2000US-0246476	
PR	08-NOV-2000	2000US-0246477	
PR	08-NOV-2000	2000US-0246478	
PR	08-NOV-2000	2000US-0246523	
PR	08-NOV-2000	2000US-0246524	
PR	08-NOV-2000	2000US-0246525	
PR	08-NOV-2000	2000US-0246526	
PR	08-NOV-2000	2000US-0246611	
PR	17-NOV-2000	2000US-0249912	
PR	17-NOV-2000	2000US-0249913	
PR	17-NOV-2000	2000US-0249914	
PR	17-NOV-2000	2000US-0249915	
PR	17-NOV-2000	2000US-0249916	

PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-483426/52.	
DR		
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
XX	metastasis -	
XX		
PS	Disclosure; SEQ ID NO 38977; 3071bp + Sequence Listing; English.	
CC	AAM54951 to AAK64702 encode the human immune/haemtopoietic antigen (I)	
CC	amino acid sequences given in AAM2170 to AAM91921. (I) have cytosolic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoitic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK67694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM62169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
XX		
Seq	Sequence 37437 BP; 10357 A; 6715 C; 6866 G; 13499 T; 0 other:	
Query Match	3.6%; Score 81.8; DB 22; Length 37437;	
Best Local Similarity	62.9%; Pred.No.1.7e-12;	
Matches 144; Conservative	0; Mismatches 82; Indels 3; Gaps 14	
OY	actctcctcttcttcctacatccacgtgagcgtaggcagcatatcctcgtcccaccagtca	866
DB		
32330	ACTGCACCTTCATCTTTTCAATGCTCATCTTCTTCCAGTATATATTCTTAAMAACAACATTA	32271
OY	867 ca----ttcctaaccgglttccaatgacctggaaccagggttgggacaacctctcgtataa	923
DB		
32270	AAATTGTCTTGCGAATGGTCGATATTCMACACAGGGGTGACGAACATTTTTCGAAAA	32211
OY	gaaaccaagaagaactattagagctctgtgycacatatgysctcagtaacaactacta	983
DB		
32210	GAGCCAGATGCGATAATATTATTAGGCTTTTGSGGCCGACGATTTCTGTTGGCAACTATTCA	32151
OY	tctctgccctctgtlgacagaaagcaatltagcaacaatalgtcaacaac	1032

Db 32150 TTCTGCTAGTACAAAGCGCTATTAACATAAGCAAC 32102  
RESULT 3  
AAK84166/C  
ID AAK84166 standard; DNA: 37442 BP.  
XX  
AC AAK84166;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38978.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001MO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198173.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 14-SEP-2000; 2000US-0233066.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236329.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000: 2000US-0249299.  
 PR 17-NOV-2000: 2000US-0249300.  
 PR 01-DEC-2000: 2000US-0250160.  
 PR 01-DEC-2000: 2000US-0250391.  
 PR 05-DEC-2000: 2000US-0251030.  
 PR 05-DEC-2000: 2000US-0251988.  
 PR 05-DEC-2000: 2000US-0256719.  
 PR 06-DEC-2000: 2000US-0251479.  
 PR 08-DEC-2000: 2000US-0251856.  
 PR 08-DEC-2000: 2000US-0251868.  
 PR 08-DEC-2000: 2000US-0251869.  
 PR 08-DEC-2000: 2000US-0251989.  
 PR 11-DEC-2000: 2000US-0251990.  
 PR 05-JAN-2001: 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-483426/52.  
 DR  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS  
 PS  
 XX  
 CC Disclosure: SEQ ID NO 38978: 3071bp + Sequence Listing; English.  
 CC  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I) `and  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastasis of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC  
 XX  
 SQ Sequence 37442 BP; 10361 A; 6715 C; 6860 G; 13506 T; 0 other;  
 Query Match 3.6%; Score 81.4; DB 22; Length 37442;  
 Best Local Similarity 63.0%; Pred. No.1.2e-12;  
 Matches 143; Conservative 0; Mismatches 81; Indels 3; Gaps 1  
 QY 809 tctctcttcttcattcatcattccatcagctgagctgagcagatcatcctgltcccccagctcaca 868  
 Db 32329 TCCACTTATTCATTCTTTCATGATGCTACTTCTTTCAGTATATATATCTCTTAAAAACAATAAAA 32270  
 QY 869 ---ttccctaccctccgttccccaatgctcgtgagccagtggtgggaacctctctctgtaaga 925  
 Db 32269 ATTGTCCTTTGCAATGTCAGATATCTAACACACAGGGGCTCAGCAACTTTTCTGAAAAA 32210  
 QY 926 accaagcaggaactattttagctctgtgtgcatatggtcccaatgctcagatcaactactcaca 985  
 Db 32209 GCCAGATGTAATATTTTAAAGCTTTGTGGGCGACAGCATTCCTGTGTGCAACTATTTCATT 32150  
 QY 986 tctgctctctgtagcagaaagcaalttagcaacaatatgtcacaac 1032  
 Db 32149 TCTGCTGCTATAGTACAAAAGCAGCTATTAACAAATTAAGTAACCAAC 32103  
 RESULT 4  
 AAH14600/c  
 ID AAH14600 standard; cDNA: 5294 BP.

[illegible]



Db 1855 GGCCATATGGTCTCTGGCAATFACCTACTCTGGCAATCAGACAGAAAGCAGCTAC 1796  
Cy 1014 caacacatgcaacaa 1030  
||| ||| |||  
Db 1795 AGACATATGTACGCA 1779  
"  
RESULT 5  
AAL36313/c  
ID AAL36313 standard; DNA; 26591 BP.  
XX  
AC AAL36313;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilicer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217996.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220863.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228824.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246603.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX MPI: 2001-451937/48.  
XX  
XX Isolated polypeptide for treating, preventing and/or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX  
PS Example 2; SEQ ID NO 2678; 781pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ALU34669-ALU37666) and proteins  
CC (AB03087-AB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pcf\_sequences.  
XX  
XX  
SQ Sequence 26591 BP; 7489 A; 6354 C; 5953 G; 6795 T; 0 other;

Query Match 3.4%; Score 77.6; DB 22; Length 26591;  
Best Local Similarity 70.3%; Pred. No. 2,4e-11;  
Matches 104; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 894 gaccacgggtggcaaaccttcctctaaagaacagacaggaactatttaggcctgt 953  
DB 23936 GACGAGGGGCTTACCAAAATTTTCTGTGAGTATAGTAAATATTTTAACTCTGC 23877  
QY 954 gtgcataatgtctcagtcacaactactactctcgtcctctgtgacgaagaacattag 1013  
DB 23876 AGGTCTGTGCTGTCTGTCTGACGTACTCACTGCATTCGATTCATGACGACCAATTGC 23817  
QY 1014 caacataatgtcaacaacaacataatgtac 1041

DB 23816 AGAAATATGAAACAAATGACGTGAC 23789  
RESULT 6  
AAK74761  
ID AAK74761 standard; DNA; 481 BP.  
XX  
XX AAK74761;  
DT 07-NOV-2001 (first entry)  
XX  
XX Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:29573.  
DE Human: Immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KM cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157182-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209486.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230439.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.



DE Human P1345 protein encoding cDNA SEQ ID NO:19/21.  
 XX Human: P1345; cancer suppression; ss.  
 KM Homo sapiens.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FT CDS 174..566  
 FT /\*tag= a  
 FT /product= "P1345"  
 XX  
 PN CN131315-A.  
 XX  
 PD 19-SEP-2001.  
 XX  
 PF 13-MAR-2000; 2000CN-0111989.  
 XX  
 PR 13-MAR-2000; 2000CN-0111989.  
 XX  
 PA (SHAN-) SHANGHAI INST ONCOLOGY.  
 XX  
 PI Gu J, Yang S;  
 XX  
 DR WPI: 2002-042193/06.  
 DR P-PSDB: ABB04713.  
 XX  
 PT New human protein able to suppress growth of cancer cells and its  
 PT encoding polynucleotide sequence -  
 PS  
 PS Claim 5; Page 33 (Disclosure): 42pp; Chinese.  
 XX  
 CC The present sequence encodes human P1345 protein, which has cancer-  
 CC suppressing activity. The present invention also describes a method  
 CC for the preparation of the protein by recombination, and the application  
 CC of the protein in treating diseases such as cancer.  
 CC  
 XX  
 SQ Sequence 1837 BP; 478 A; 471 C; 407 G; 481 T; 0 other;

Query Match 3.28; Score 73; DB 24; Length 1837;  
 Best Local Similarity 64.38; Pred. No. 1.2e-10;  
 Matches 109; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
 Oy 918 tgaagaagaccagagaaactatttaggtctgtgctcattatgctcagtcacac 977.  
 Db 1639 tttaaaagaaacagagacaaatatttagcttctgtgacacatgctcgtcacac 1698  
 Oy 978 tactatctctgctctctgtagcagaaagcaattgcaacaatattgtaacaacattatg 1037  
 Db 1699 tactcagctctgtattgttaacaaagcagctgtgacagatgtaataatgaaatg 1758  
 Oy 1038 tgaaccatgaaacatttattatcatgatacgagaacctgaaataa 1086  
 Db 1759 tggctgtgttccataataaacttatttcccaaaaaaaaaaaaaa 1807

RESULT 8  
 AAH98619/c  
 ID AAH98619 standard; cDNA; 628 BP.  
 XX  
 AC AAH98619;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Human EST-derived coding sequence SEQ ID NO: 476.  
 XX  
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KM diagnostics; forensic test; gene mapping; genetic disorder;  
 KW biodiversity; gene therapy; nutrition; ss.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200154477-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02687.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 PR 17-JUL-2000; 2000US-0617746.  
 PR 03-AUG-2000; 2000US-0631451.  
 PR 15-SEP-2000; 2000US-0663870.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 XX  
 DR WPI: 2001-476164/51.  
 DR P-PSDB: AAM23960.  
 XX  
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use -  
 PS  
 PS Claim 1; Page 514; 1275pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
 CC of the invention.  
 XX  
 SQ Sequence 628 BP; 190 A; 122 C; 131 G; 185 T; 0 other;

Query Match 3.18; Score 71.6; DB 22; Length 628;  
 Best Local Similarity 60.98; Pred. No. 1.7e-10;  
 Matches 151; Conservative 0; Mismatches 94; Indels 3; Gaps 2;  
 Oy 894 gaccagcttggaacactcttcctgtaagaacagacagaaactatttagctctgt 953  
 Db 452 GACCAGGCTTCAGACAACTTTTCTGTAAAGGCAAAATGTAATTTTAGCTTGT 393  
 Oy 954 gtgcacatgctctcagcaacactactcactctctgtgacgagaaacattag 1013  
 Db 392 GGACCAATATGCTTC--TGCCTAACTACTCAATTCCTGCCATTTGAAGTCAACACCAAT 335  
 Oy 1014 caacaatatgtcaacaacaacatatgtgaccccaatgaactatttattatgatacga 1073  
 Db 334 GTA-AATGAATGTCAGTCGCTGTATTATTGTGACACAGTAATTAATTTCAATAT 276  
 Oy 1074 aacctgaanaataatgctctcttcttgaatttctcccaatcaatgaacagcaaaact 1133  
 Db 275 TTCACATTAATTAATTAATTTTGTGATTTTTCCTCCACAAATTMAAAATTTTAAAGCC 216  
 Oy 1134 actctttag 1141  
 Db 215 ATGCTTAG 208

RESULT 9  
 AAA41353/c  
 ID AAA41353 standard; cDNA; 286 BP.  
 XX  
 AC AAA41353;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Human secreted expressed sequence tag SEQ ID NO:93.  
 XX  
 XX Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW

KW		immunomodulatory; haematopoietic; chemokinetic; analgesic; hemostatic;
KW		antiviral; antinflammatory; cytosolic; antibacterial; antifungal;
KM		thrallalytic; antididiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW		antiulcer; osteopathic; neuroprotective; nootropic; antisporiatric;
KV		cerebroprotective; anticoagulant; antidepressant; gene therapy;
KW		vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KV		insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW		lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW		central nervous system disorder; Alzheimer's disease; stroke;
KM		Parkinson's disease; Huntington's disease; coagulation disorder;
KW		haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KX		tumour; infection; depression; psoriasis; ss.
OS	Homo sapiens.	
XX		
PN	WO200021990-AI.	
PD	20-APR-2000.	
PE	15-OCT-1999; 99MO-US24205.	
PR	15-OCT-1998; 98US-0104435.	
XX	(GENET ) GENETICS INST INC.	
PA	Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C;	
PI	Mertberg D, Treacy M;	
DR	MPJ, 2000-317937/27.	
XX		
PT	Isolated polynucleotides, and encoded proteins, comprising secreted	
PT	expressed sequence tags (seSTS), useful for treating various disorders	
PS	such as autoimmune, infectious, and central nervous system disorders -	
Claim 1; Page 198: 618pp; English.		
xx		
CC	AAA41261 to AAA41419 represent specifically claimed secreted expressed	
CC	sequence tags (seSTS), isolated from human, mouse, xenopus and rat	
CC	tissue sources. The seSTS can have a range of activities depending on	
CC	the tissues they were isolated from. The activities include:	
CC	chemotactic; proliferative; immunomodulatory; haematopoeitic;	
CC	chondrokinetic; analgesic; haemostatic; thrombolytic; antinflammatory;	
CC	cytostatic; antibacterial; antifungal; antiviral; antidiabetic;	
CC	antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;	
CC	nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;	
CC	anticovulsant; and antidepressant. The seSTS can be used for gene	
CC	therapy and in vaccines. The seSTS are useful as probes for the	
CC	identification and isolation of full-length cDNAs and genomic DNA	
CC	molecules which correspond to the seSTS. Proteins encoded by the seSTS	
CC	are useful in assays for determining biological activity and raising	
CC	antibodies. They may be useful for treatment of autoimmune disorders	
CC	(multiple sclerosis, insulin dependent diabetes), allergic conditions	
CC	(asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,	
CC	osteoporosis, osteoarthritis, central nervous system disorders	
CC	(Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation	
CC	disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's	
CC	disease), tumours, bacterial, fungal or viral infections, depression and	
CC	psoriasis. AAA41320 to AAA41345 represent linker variants which are given	
CC	in the exemplification of the present invention.	
SQ		
XO	Sequence 286 BP; 88 A; 63 C; 50 G; 85 T; 0 other;	
Query Match	3.1%; Score 71.4; DB 21; Length 286;	
Best Local Similarity	72.1%; Pred.No.1.2e-10;	
Matches	93; Conservative 0; Mismatches 36; Indels 0; Gaps 0;	
DY	899 aggttgggcaaaccttcctgtaagaacaacgaaggaaacttttagcctctgtgcgcc 958	
Db	149 AGGCACGATTAACCTTTTCTGTAAAGACGAACAAGTAATATTITTAGGCCGTGGTGTC 90	
OY	atattgctccagtcaacatactcatctctgcctctgttagcacgaagaaccttagcaaca 1018	

Db	89	ATATGCTCTCTGCGCACTACACAACTCTGCTTCTTGTATAAAGACGATAGACA	30
Qy	1019	atatgtcaaa 1027	
Db	29	ATATGTAAA 21	
RESULT 10			
ID	AAc27788/c		
XX	AAc27788 standard; cDNA; 407 BP.		
XX			
AC	AAc27788;		
XX			
DT	06-OCT-2000 (first entry)		
XX			
DE	Human secreted protein 5' EST, SEQ ID NO: 31863.		
XX			
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;		
KW	gene therapy; chromosome mapping; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EPI033401-A2.		
PD	06-SEP-2000.		
XX			
PF	21-FEB-2000; 2000EP-0200610.		
XX			
PR	26-FEB-1999; 990S-0122487.		
XX			
PA	(GEST ) GENSET.		
PI	Dumas Mline Edwards J, Duclert A, Giordano J;		
XX			
DR	WPI: 2000-500381/45.		
XX			
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for		
XX	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for		
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -		
XX			
PS	Claim 1; SEQ ID 31863; 71pp + CD-ROM; English.		
XX			
CC	The present sequence is one of a large number of 5' ESTs derived from		
CC	mRNAs encoding secreted proteins. No ORF has yet been conclusively		
CC	identified within the present sequence. The 5' ESTs were prepared from		
CC	total human RNAs or polyA+ RNAs derived from 30 different tissues. EST		
CC	sequences usually correspond mainly to the 3' untranslated region (UTR)		
CC	of the mRNA because they are often obtained from oligo-dT primed cDNA		
CC	libraries. Such ESTs are not well suited for isolating cDNA sequences		
CC	derived from the 5' ends of mRNAs and even in those cases where longer		
CC	cDNA sequences have been obtained, the full 5' UTR is rarely included.		
CC	5' ESTs are derived from mRNAs with intact 5' ends and can therefore be		
CC	used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used		
CC	in diagnostic, forensic, gene therapy and chromosome mapping procedures.		
CC	They are used to obtain upstream regulatory sequences and to design		
CC	expression and secretion vectors.		
XX			
SO	Sequence 407 BP; 118 A; 72 C; 84 G; 127 T; 6 other;		
Query Match 3.1%; Score 71.2; DB 21; Length 407;			
Best Local Similarity 63.6%; Pred. NO.1.7e-10;			
Matches 124; Conservative 0; Mismatches 69; Indels 2; Gaps 1;			
Qy	887	atgacctgagccaggattgggcaaacctcttcctgtaaaagacagacagagattatllag 946	
Db	402	ATGCCTAGACACACAGACGGCAAAATTTTCTGTAAAGAGCCAGACAGTAATATGTTAG 343	
Qy	947	gtctcgtgtgcacatatggtctcaatgcaaacactacatctctgctctcgtagcacgaaag 1006	
Db	342	GGTTTGCAGGGCTACATGATCTCTGTATTGACAACTACACATGACGCTCCCAATGACAAAAG 283	
Qy	1007	caattagcaacacatatgtcaaacacacatatgtgaccccatgaaacattatattatattat 1056	



QY 1009 atgagcaacatgtcacaaca 1031  
 DB 8849 AGTATGACAAATATATMAACAA 8827

## RESULT 13

AAH99747  
 ID AAH99747 standard; CDNA; 8029 BP.  
 AC AAH99747;  
 XX  
 XX 16-OCT-2001 (first entry)  
 DE Human protein encoding CDNA sequence SEQ ID NO:582.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KW antibacterial; endocrine; cardiac; central nervous system; vitreous;  
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;  
 KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;  
 KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;  
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder; ss.

XX Homo sapiens.  
 OS  
 PN W0200153455-A2.  
 PD 26-JUL-2001.

XX 22-DEC-2000; 2000MO-US35017.  
 PF  
 XX 23-DEC-1999; 99US-0471275.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 XX

PA (HYSE-) HYSKO INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457603/49.

DR P-PSDB; AAM25806.

PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

PS Claim 1: Page 622-624; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAH25963. The proteins can have activities based on the tissues and  
 CC cells they are expressed in, such as: antinflammatory; antirheumatic;  
 CC antitumor; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; vitreous; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;  
 CC anticancer; osteopathic; dermatological; antiallergic; antisthmatic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.

XX Sequence 8029 BP; 2064 A; 1942 C; 1905 G; 2118 T; 0 other;

Query Match 3.0%; Score 69.4; DB 22; Length 8029;  
 Best Local Similarity 57.2%; Pred. NO. 3.2e-09;  
 Matches 202; Conservative 0; Mismatches 121; Indels 30; Gaps 3;

QY 889 gcttgaccacggttgggaacactctctgttaagaacacagagacattttaggc 948  
 DB 682 gtctaagatcagagcgtcgcaacatttcttaagagagacagtgtaactttaggc 741  
 QY 949 tctgtgtccatgtgtctcagtcacactactctctgctcttagaac----- 1001  
 DB 742 ttgcagcgtactactctccctgcccaactatcactctgcagagaagacgccacca 801  
 QY 1002 -----gaaagcaattagacaacatgttcaacaacacatagtagaccacatga 1048  
 DB 802 caatacacaataagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 861  
 QY 1049 aaacttattt-atagtatagcaggaacacgtgaataatgtcttctttagattttc 1107  
 DB 862 aactgtatctcattagatatttcaagtgcatgaataatgtcttctttagattttc 921  
 QY 1108 cccaatcattaaagaacgaacacacacacacacacacacacacacacacacacacac 1167  
 DB 922 cccaatcattgaac 972  
 QY 1168 tagcagtgagcagcgtgattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1220  
 DB 973 aacagtgagcagcagcgtttagtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1025

## RESULT 14

AAH99306/C  
 ID AAH99306 standard; DNA; 46765 BP.

XX AAH99306;

XX 12-MAR-2002 (first entry)

DE DNA encoding Aldehyde dehydrogenase 5 family, member A1 (ALDH5A1).

XX Aldehyde dehydrogenase 5 family member A1; ALDH5A1;  
 KW succinate-semialdehyde dehydrogenase; gene therapy; probe;  
 KW antisense technology; allele specific oligonucleotide; ASO;  
 KW 4-hydroxybutyric aciduria; metabolic disease; transgenic animal;  
 KW chromosome 6p22; ds.

XX Homo sapiens.

OS Homo sapiens.

PN replace(3749,T)

PT variation

PT variation

PT variation

PT variation

PT variation

PT variation

PT variation  
 /tag- e  
 /standard\_name-  
 /tag- f  
 /standard\_name-  
 /standard\_name- "Single nucleotide polymorphism"

```

FT      exon      4050..4403
FT      /*tag= g
FT      /number= 1
FT      4050..42765
FT      /*tag= h
FT      /product= "ALDH5A1"
FT      /note= "Aldehyde dehydrogenase 5 family member A1.
FT      Specifically claimed in claim 25"
FT      replace(4059,G)
FT      /*tag= i
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4111,A)
FT      /*tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4155,C)
FT      /*tag= k
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4179,G)
FT      /*tag= l
FT      /standard_name= "Single nucleotide polymorphism"
FT      4404..11575
FT      /*tag= m
FT      /number= 1
FT      replace(4430,G)
FT      /*tag= n
FT      /standard_name= "Single nucleotide polymorphism"
FT      11576..11659
FT      /*tag= o
FT      /number= 2
FT      11660..12315
FT      /*tag= p
FT      /number= 2
FT      12316..12486
FT      /*tag= q
FT      /number= 3
FT      replace(12445,T)
FT      /*tag= r
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(12442,T)
FT      /*tag= s
FT      /standard_name= "Single nucleotide polymorphism"
FT      12487..13921
FT      /*tag= t
FT      /number= 3
FT      replace(12559,T)
FT      /*tag= u
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(13855,G)
FT      /*tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(13917,T)
FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      13922..14038
FT      /*tag= x
FT      /number= 4
FT      replace(14021,T)
FT      /*tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      14039..24219
FT      /*tag= z
FT      /number= 4
FT      replace(14046,A)
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(14179,A)
FT      /*tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      24220..24363
FT      /*tag= ac
FT      /number= 5
FT      24364..29453
FT      /*tag= ad

```

```

FT      intron    24598..31819
FT      /*tag= ae
FT      /number= 6
FT      replace(29397,T)
FT      /*tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      29454..29597
FT      /*tag= ag
FT      /number= 6
FT      replace(29469,A)
FT      /*tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      31820..31978
FT      /*tag= ai
FT      /number= 7
FT      31979..37049
FT      /*tag= aj
FT      /number= 7
FT      37050..37219
FT      /*tag= ak
FT      /number= 8
FT      37220..41171
FT      /*tag= al
FT      /number= 8
FT      41172..41230
FT      /*tag= am
FT      /number= 9
FT      replace(41217,C)
FT      /*tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      41231..42559
FT      /*tag= ao
FT      /number= 9
FT      replace(41237,A)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(41403,C)
FT      /*tag= aq
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(42380,T)
FT      /*tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      42560..42765
FT      /*tag= as
FT      /number= 10
FT      exon      WO200190119-A2.
FT      PN      29-NOV-2001.
FT      PD      21-MAY-2001; 2001WO-US16558.
FT      PP      19-MAY-2000; 2000US-205849P.
FT      PR      (GENA-) GENAISSANCE PHARM INC.
FT      PA      Kilem SE, Koshy B, Tanguay DA.
FT      PI      WPI: 2002-089912/12.
FT      DR      P-PSDB: AA073594.
FT      XX      New genetic variants of human aldehyde dehydrogenase 5 family, member
FT      PT      AL, ALDH5A1 gene for treating metabolic diseases and for expressing
FT      PT      ALDH5A1 protein useful in identifying drugs to treat 4-hydroxybutyric
FT      PT      aciduria
FT      PS      Disclosure: Fig 1; 151pp; English.
FT      XX      The invention describes an isolated polynucleotide comprising a
FT      CC      nucleotide sequence which is a polymorphic variant of a reference
FT      CC      sequence for the aldehyde dehydrogenase 5 family, member A1
FT      CC      (succinate-semialdehyde dehydrogenase) (ALDH5A1) gene or its fragment.

```



CC The polypeptide is useful for screening for drugs targeting it by  
CC contacting the ALDH5A1 polymorphic variant with a candidate agent and  
CC assaying for binding activity. The polypeptide and haplotypes are useful  
CC for identifying an association between a trait such as a clinical  
CC response to a drug targeting ALDH5A1 and a haplotype ALDH5A1 gene.  
CC Transgenic animals are also useful for studying expression of the ALDH5A1  
CC isogenes in vivo, for in vivo screening and testing of drugs against  
CC ALDH5A1 protein and for testing the efficacy of therapeutic agents and  
CC compounds for 4-hydroxybutyric aciduria and metabolic diseases in a  
CC biological system. Antibodies are useful for diagnostic and prognostic  
CC formats and therapeutic methods, for immunoprecipitating the polypeptide  
CC from solution, for detecting ALDH5A1 protein isoforms in biological  
CC samples, frozen tissue sections, for use in immunocytochemical,  
CC immunohistochemical and immunofluorescence techniques. The polynucleotide  
CC is useful for gene therapy and antisense gene therapy. This sequence  
CC encodes ALDH5A1 (located on chromosome 6p22), and forms a reference  
CC sequence on which polymorphic sequences encoding isoforms of the ALDH5A1  
CC protein are based, described in the method of the invention.  
XX

Query Match 3.0%; Score 69.4; DB 24; Length 46765;  
Best Local Similarity 57.2%; Pred. No. 8.8e-09;  
Matches 202; Conservative 0; Mismatches 121; Indels 30; Gaps 3;

OY 889 gccggaccaggttgggcaaacctcttcgttaaagacagacaggaactatttaggc 948  
DB 45569 GCTTAGATCAGAGGAGCTGGCAACTTTTCTTAAAGAGCAAACTAGTAATCTTTAGGC 45530  
OY 949 tctgtgcatatgcatgctcagtcacaactactcactcctcctcgttagac----- 1001  
DB 45529 TTTGAGGCTACTGCTCCCTGCCAATTCACCTGCGCAAGAACACTGCCAGA 45470  
OY 1002 -----gaaagcaattggcaacaatltgtcaacaacataatgaccacatga 1048  
DB 45469 CAAATCAATTAATGATGGGTGGCTATGCTTAATTAACATTTTGTATTGACACTGA 45410  
OY 1049 aaacttattt-attatgatacggaaacctgaataatagtcctcttcttgaatttttc 1107  
DB 45409 AATCGAATTCATGATGAAATTTTCAGGGTGCATGAAATATTGTTCTCTTTGATTTCG 45350  
OY 1108 cccaatcattaaaaagtaaaactactctttagtctgcaaggttaagccattcctcagct 1167  
DB 45349 CCCAATCATTTGAAAACATTAATCAATCTTACCTCAGG-----TTGTATAA 45299  
OY 1168 tagcagtgagcagctgatttggccttgcactacagtcgccaatccctgat 1220  
DB 45298 AACAGGTGGCAGGCCAGGTTTGATTCGTGGGCCATAGTTTGCACACCTGAT 45246

RESULT 15  
AAK64813/C  
ID AAK64813 standard; DNA; 47319 BP.  
XX  
AC AAK64813;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19625.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO200157182-AZ.  
XX  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX  
XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.

02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0244675.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251989.  
PR 06-DEC-2000; 2000US-0251719.  
PR 06-DEC-2000; 2000US-0251719.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251858.  
PR 08-DEC-2000; 2000US-0251859.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 19625; 3071pp + Sequence Listing; English.  
XX

CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins, and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX

-SQ Sequence 47319 BP; 13708 A; 8453 C; 9720 G; 15438 T; 0 other;

Query Match 3.0%; Score 69.4; DB 22; Length 47319;  
Best Local Similarity 66.2%; Pred. No. 8.9e-09;  
Matches 100; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 881 ttcccatgctgagccaggttggaactctctggaagacagaggaacta 940  
DB 18549 TTTACTAGTCTAAATGACAGATGGCAACTTTTCTGCAAGACCGCATTAATTA 18490  
OY 941 ttttagctctgctgcatatgltcagcaactactactctgctctgtaga 1000  
DB 18489 TTTAGGCTTGTAAAGCCACAGCTCTCTACGGCAACCACTCACTCATATAGCC 18430  
OY 1001 cgaagcaattagcaacatatgtccaacaa 1031  
DB 18429 CCAAGCTCTACAGATGATATTAATTA 18399

Search completed: August 16, 2002, 05:33:21  
Job time: 10905 sec

• Mon Aug 19 06:46:33 2002

us-09-515-369b-1.rng

Page 17



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 03:40:03 : Search time 5721.79 Seconds  
(without alignments)  
8643.690 Million cell updates/sec

Title: US-09-515-369B-1

Perfect score: 2286  
1 taatagcactcatataggg.....tgacttcacgctcggagc 2286

Scoring table: IDENTITY,NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending-Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq:\*  
17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:\*  
20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq:\*  
21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq:\*  
22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq:\*  
23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq:\*  
24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq:\*  
25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:\*  
26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:\*  
27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:\*  
28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:\*  
29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq:\*  
30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq:\*  
31: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*  
32: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*  
33: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*  
34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*  
35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*  
36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*  
37: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq:\*  
38: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq:\*  
39: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq:\*  
40: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq:\*  
41: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq:\*  
42: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq:\*  
43: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq:\*

SUMMARIES					Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
Result No.	Score	Query Match	Length DB	ID	Description	
1	2286	100.0	2286	19	US-09-515-369B-1	
2	420	18.4	426	19	US-09-528-409-108247	
3	420	18.4	426	35	US-09-933-524-108247	
4	420	18.4	426	35	US-09-933-524A-108247	
5	93	4.1	338	22	PCT-US01-15674A-421	
6	93	4.1	338	22	US-09-573-080A-421	
7	87	3.8	196093	63	US-60-245-228-200	
8	87	3.8	196093	64	US-60-258-272-80	
9	86	3.8	224	1	PCT-US01-15674A-420	
10	86	3.8	224	22	US-09-573-080A-420	
11	86	3.8	76534	63	US-60-243-468-358	
12	83.2	3.6	179040	61	US-60-226-176-1401	
13	83.2	3.6	179040	62	US-60-233-468-1401	
14	83.2	3.6	179040	70	US-60-313-371-1401	
15	83	3.6	163174	60	US-60-212-664-256	
16	82.8	3.6	1007	24	US-09-634-306B-119945	
17	82.8	3.6	1007	24	US-09-634-306B-119946	
18	82.4	3.6	17947	60	US-60-212-358-69	
19	81.8	3.6	288	14	US-09-070-694-1633	
20	81.8	3.6	288	43	US-60-044-030-1633	
21	81.8	3.6	511	56	US-60-172-360-1441	
22	81.8	3.6	511	57	US-60-185-216-346	
23	81.8	3.6	37437	1	PCT-US01-01354-38977	
24	81.8	3.6	37437	30	US-09-764-905-38977	
25	81.8	3.6	37437	35	US-09-950-083-9956	
26	81.8	3.6	37437	37	US-10-092-399-38977	
27	81.4	3.6	37442	1	PCT-US01-01354-38978	
28	81.4	3.6	37442	30	US-09-764-905-38978	
29	81.4	3.6	37442	35	US-09-950-083-9957	
30	81.4	3.6	37442	37	US-10-092-399-38978	
31	81.2	3.6	161395	61	US-60-226-176-1603	

```

c 32 81.2 3.6 161395 62 US-60-233-468-1603 Sequence 1603, Ap
c 33 81.2 3.6 161395 70 US-60-313-371-1603 Sequence 1603, Ap
c 34 80.8 3.5 32768 59 US-60-208-965-68 Sequence 68, Appl
c 35 79.4 3.5 503 16 US-09-235-076-12893 Sequence 12893, A
c 36 79.4 3.5 503 16 US-09-248-797-14353 Sequence 14353, A
c 37 79.4 3.5 503 17 US-09-332-782-12893 Sequence 12893, A
c 38 79.4 3.5 503 29 US-09-737-223-12893 Sequence 12893, A
c 39 79.4 3.5 503 34 US-09-925-564-34353 Sequence 34353, A
c 40 79 79 749 24 US-09-634-306B-151080 Sequence 151080,
c 41 79 3.5 749 24 US-09-634-306B-151081 Sequence 151081,
c 42 78.4 3.4 312283 1 PCT-US00-27620-91 Sequence 91, Appl
c 43 78.2 3.4 7748 24 US-09-634-306B-7748 Sequence 7748, Ap
c 44 78.2 3.4 118794 63 US-60-245-468-310 Sequence 310, Appl
c 45 78.2 3.4 118794 63 US-60-245-742-53 Sequence 53, Appl
```

## ALIGNMENTS

```

RESULT 1
; Sequence 1, Application US/09515369B
; GENERAL INFORMATION:
; APPLICANT: Paul B. Fisher
; APPLICANT: Malavi T. Medireddi
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; FILE REFERENCE: 34611 070050.1685
; CURRENT APPLICATION NUMBER: US/09/515, 369B
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Human
; US-09-515-369b-1
```

```

Query Match 100.0%; Score 2286; DB 19; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 taataagactacataatgagcgtcgactcgatcaacttttgaaacaaagctcgctgcct 60
D 1 taataagactacataatgagcgtcgactcgatcaacttttgaaacaaagctcgctgcct 60
QY 61 ccaaaagctgtactcaatactagatctcaactgactgtcgagcgaaggtctcctaggtct 120
D 61 ccaaaagctgtactcaatactagatctcaactgactgtcgagcgaaggtctcctaggtct 120
QY 121 ctctctgacctctctctgaaagtaataagctatcgaagctcgaagctcgagcc 180
D 121 ctctctgacctctctctgaaagtaataagctatcgaagctcgaagctcgagcc 180
QY 181 aggcacatgttgcgtgaactatccatgtatatagtatctctccctccacagaagtgagct 240
D 181 aggcacatgttgcgtgaactatccatgtatatagtatctctccctccacagaagtgagct 240
QY 241 actcaagatcccaagtgtaaccctgagcgaagcgaagtgatlatccatgacctatgctctc 300
D 241 actcaagatcccaagtgtaaccctgagcgaagcgaagtgatlatccatgacctatgctctc 300
QY 301 gtccagagctgaccttaacagctatccacactgctccctcccgctcctatcgagaa 360
D 301 gtccagagctgaccttaacagctatccacactgctccctcccgctcctatcgagaa 360
QY 361 cagtagtctaggaattcagctgcccctgagggtcattctccctcctcagcttctcctta 420
D 361 cagtagtctaggaattcagctgcccctgagggtcattctccctcctcagcttctcctta 420
QY 421 gctgtctcctgctccacactcactatctactcagacactcactcagcttctctctctg 480
D 421 gctgtctcctgctccacactcactatctactcagacactcactcagcttctctctctg 480
```

```

D 421 gctgtctcctgctccacactcactatctactcagacactcactcagcttctctctctg 480
QY 481 tctcaatcctgctctcttgacaattcttccatagtagtaagtaagggtctcttgat 540
D 481 tctcaatcctgctctcttgacaattcttccatagtagtaagtaagggtctcttgat 540
QY 541 gccctaaatcccaatcgtgaggaaggggagtgagggaagagtgagctgtgaggtctg 600
D 541 gccctaaatcccaatcgtgaggaaggggagtgagggaagagtgagctgtgaggtctg 600
QY 601 gctctctctgagagtgagagctgagcctcccgaggaagaaagatctcaggtctgagaa 660
D 601 gctctctctgagagtgagagctgagcctcccgaggaagaaagatctcaggtctgagaa 660
QY 661 gctatagcgaagcagacgctgagcgaagagatgcaaaagagatcttctgtctctaaaga 720
D 661 gctatagcgaagcagacgctgagcgaagagatgcaaaagagatcttctgtctctaaaga 720
QY 721 aatlaaacaactgagtagatgagagggaggggagtggtgtgagcagaagatgggaa 780
D 721 aatlaaacaactgagtagatgagagggaggggagtggtgtgagcagaagatgggaa 780
QY 781 gactgtgcaagagtgtagtctactcaatcctcctctctctcctcctcctcctcagtgag 840
D 781 gactgtgcaagagtgtagtctactcaatcctcctctctctcctcctcctcctcagtgag 840
QY 841 aggcagatctatcctgtcccccaggtcacatctcctcctcctcctcctcctcctcctcct 900
D 841 aggcagatctatcctgtcccccaggtcacatctcctcctcctcctcctcctcctcctcct 900
QY 901 gctggcgaacactctccgtgtaaaagacagaagaaactatcttaagcctctgtgagcat 960
D 901 gctggcgaacactctccgtgtaaaagacagaagaaactatcttaagcctctgtgagcat 960
QY 961 atggtctcagtagcaacacacacacacacacacacacacacacacacacacacacacac 1020
D 961 atggtctcagtagcaacacacacacacacacacacacacacacacacacacacacacac 1020
QY 1021 atgcaacaacaataatgtagaccccaatgaaactatcttaatgtagacacacacacacac 1080
D 1021 atgcaacaacaataatgtagaccccaatgaaactatcttaatgtagacacacacacacac 1080
QY 1081 aaataatgctctctctctctctctctctctctctctctctctctctctctctctctct 1140
D 1081 aaataatgctctctctctctctctctctctctctctctctctctctctctctctctct 1140
QY 1141 ggtcgcaaggttaagcactctcagcttagcagtgagcaggtctgagctgtgagcact 1200
D 1141 ggtcgcaaggttaagcactctcagcttagcagtgagcaggtctgagctgtgagcact 1200
QY 1201 acaagttgcgaatccctctatctcccaaaatgattctcccaagagtggtgcaaatctat 1260
D 1201 acaagttgcgaatccctctatctcccaaaatgattctcccaagagtggtgcaaatctat 1260
QY 1261 gggaaagtgctgtagtaaaacagaagtaagaaacacacacacacacacacacacacacac 1320
D 1261 gggaaagtgctgtagtaaaacagaagtaagaaacacacacacacacacacacacacacac 1320
QY 1321 atgcaagagctctcctaaatgacacacacacacacacacacacacacacacacacacac 1380
D 1321 atgcaagagctctcctaaatgacacacacacacacacacacacacacacacacacacac 1380
QY 1381 acaagaatctcctcaaaacacacacacacacacacacacacacacacacacacacacacac 1440
D 1381 acaagaatctcctcaaaacacacacacacacacacacacacacacacacacacacacacac 1440
QY 1441 atgcccctcactgagacacacacacacacacacacacacacacacacacacacacacac 1500
D 1441 atgcccctcactgagacacacacacacacacacacacacacacacacacacacacacac 1500
QY 1501 atgcaagacactgagacacacacacacacacacacacacacacacacacacacacacac 1560
D 1501 atgcaagacactgagacacacacacacacacacacacacacacacacacacacacacac 1560
```



Db 365 GTAAACACGCGCCCTCCAGAGAAACAGATTCAAGCTGGGGCAGCTATAGCCAGCAG 306  
Oy 676 actgctggccaggagatgcaaaaggagatcttctgttctgaagaaataaacaacactga 735  
Db 305 ACTGCTGGCCAGGAGTTTCAAAAGAGTATTTTGTGTTAAAGAAATAAACACACTGA 246  
Oy 736 gtaagatggaaggagggtgtgtgtgcaagagagatctgggaagagcttgcgaaggt 795  
Db 245 GTATGAGATGAGGAGGAGGGGTGTGTGTCCAGAGAGATTTGGGAAGACTGCCAAGG 186  
Oy 796 gtgtctactcactctctctctctctctctctcaactcagagcttgaagcagttatccgt 855  
Db 185 GTGTTCTACTCACT 126  
Oy 856 ccccaacgtacatctcactcctccgtctcccatgcttgcgaagcaggttggcaaacctct 915  
Db 125 CCCCACGTCACATCT 66  
Oy 916 cctgtgaagaaccag 975  
Db 65 CCGTAAAGAACACAGAGAACTATTTAGGCTGTGTGTCCATATGTCTCAGTCACA 6

RESULT 4  
US-09-933-524A-108247/C  
Sequence 108247, Application US/09933524A  
GENERAL INFORMATION:  
APPLICANT: Drmanac, Radoje T.  
APPLICANT: Labat, Ivan  
APPLICANT: Stache-Crain, Birgit  
APPLICANT: Dickson, Mark  
APPLICANT: Jones, Lee W.  
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained  
TITLE OF INVENTION: From Various Libraries  
FILE REFERENCE: 774  
CURRENT APPLICATION NUMBER: US/09/933\_524A  
CURRENT FILING DATE: 2001-08-20  
PRIOR APPLICATION NUMBER: 09/528,409  
PRIOR FILING DATE: 2000-05-17  
NUMBER OF SEQ ID NOS: 116231  
SOFTWARE: Hy-Patent.pl Version 3.1  
SEQ ID NO 108247  
LENGTH: 426  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-933-524A-108247

Query Match 18.4%; Score 420; DB 35; Length 426;  
Best Local Similarity 100.0%; Pred. No. 2.1e-108;  
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 556 ggtggaagaggaggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 615  
Db 425 GGTGGAGAGGGGGAGTGGGGAGAGAGCTGCGCTGTGGGGCTGCTCACTTCTGGAGG 366  
Oy 616 gtaagactggggccctccagaaagagatcaagctgtgtgcaagctataagccaagcag 675  
Db 365 GTAAAGACTGGGCGCTCCAGAGAACAAAGAGTTCAGGCTGGTGCACCTATAGCCAGCAG 306  
Oy 676 actgctggccaggagatgcaaaaggagatcttctgttctgaagaaataaacaacactga 735  
Db 305 ACTGCTGGCCAGGAGTTTCAAAAGAGTATTTTGTGTTAAAGAAATAAACACACTGA 246  
Oy 736 gtaagatggaaggagggtgtgtgtgcaagagagatctgggaagagcttgcgaaggt 795  
Db 245 GTATGAGATGAGGAGGAGGGGTGTGTGTCCAGAGAGATTGGGAAGAGTCTGCCAAGG 186  
Oy 796 gtgtctactcactctctctctctctctctctcaactcagagcttgaagcagttatccgt 855  
Db 185 GTGTTCTACTCACT 126

Oy 856 ccccaacgtacatctcactcctccgtctcccatgcttgcgaagcaggttggcaaacctct 915  
Db 125 CCCCACGTCACATCT 66  
Oy 916 cctgtgaagaaccag 975  
Db 65 CCGTAAAGAACACAGAGAACTATTTAGGCTGTGTGTCCATATGTCTCAGTCACA 6

RESULT 5  
PCT-US01-15674A-421/C  
Sequence 421, Application PC/TUS0115674A  
GENERAL INFORMATION:  
APPLICANT: JOAN, KNOLE H  
APPLICANT: ROGAN, PETER K  
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT  
FILE REFERENCE: 30307  
CURRENT APPLICATION NUMBER: PCT/US01/15674A  
CURRENT FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 613  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 421  
LENGTH: 338  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: repeat\_region  
LOCATION: (1)..(338)  
OTHER INFORMATION: cheshire\_b  
PCT-US01-15674A-421

Query Match 4.1%; Score 93; DB 1; Length 338;  
Best Local Similarity 65.8%; Pred. No. 4.7e-15;  
Matches 177; Conservative 1; Mismatches 66; Indels 25; Gaps 2;

Oy 900 gttggggaacactctctctgtgaagaaccagagagagagagagagagagagagagagag 959  
Db 335 GGGTCGGCAAACTTTTGTGTAAGGGCCAGATGTAATATTTTAGGCTTGGCGGCCA 276  
Oy 960 tatgtcccaagtcacaactactcctctcctctgtgagcagaaagcaatlagcaacaa 1019  
Db 275 TAGGTCCTCTGTGCGCAACTACTCAACTCTGCGGTGTAGCCCAAAAGCAGCATAGACAA 216  
Oy 1020 tatgtcaacaacatataatgagcccat--gaaacttattatattatagagagaaacc 1077  
Db 215 TATATTAACGATGGCGT 156  
Oy 1078 tga-----aaatagtctctctctctctctctctctcccaatc 1114  
Db 155 TGAATTTTCAATTAATTTTCAATGTCACAAATATATCTTCTTTGATTTTCTTCCAAAC 96  
Oy 1115 atlaaaacgtlaaaactactctaggt 1143  
Db 95 ATTTAAATGTAAATTAACCAATCTTAGTT 67

RESULT 6  
US-09-573-080A-421/C  
Sequence 421, Application US/09573080A  
GENERAL INFORMATION:  
APPLICANT: JOAN, KNOLE  
APPLICANT: ROGAN, PETER  
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENER  
FILE REFERENCE: 30307  
CURRENT APPLICATION NUMBER: US/09/573\_080A  
CURRENT FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 479  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 421  
LENGTH: 338  
TYPE: DNA  
ORGANISM: Homo sapiens



```

FEATURE:
NAME/KEY: repeat_region
LOCATION: (1)-(338)
OTHER INFORMATION: cheshire_b
PUBLICATION INFORMATION:
AUTHORS: Jurka, J; Malchuk, J; Miosavljic, A
TITLE: Prototypic sequences for human repetitive DNA
JOURNAL: Journal of Molecular Evolution
VOLUME: 35
ISSUE: 4
PAGES: 286-291
DATE: 1992-10-
DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
DATABASE ENTRY DATE:
DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-421

```

```

Query Match 4.1%; Score 93; DB 22; Length 338;
Best Local Similarity 65.8%; Pred. No. 4.7e-15;
Matches 177; Conservative 1; Mismatches 66; Indels 25; Gaps 2;

```

```

OY 900 ggttggaacacgtcttcctgtaagaacaggaactatttagctctgtgtcca 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 GGGTCGGCAACCTTTCTGTAAAGGGCCAGATGTAATAATTTAGCTTGGGGCCA 276

OY 960 tatgtctcagtcacacactctctctctctctctctctctctctctctctct 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 TAGGGTCTCTGCCAATCTCACTGCTGCGTTGAGCCCAACAGCCATAGACAA 216

OY 1020 tatgtcaacaacacatgtgaccccat--gaaacttattattatgatacggaaac 1077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 TATATTAACGAATGGGGTGGCTGTGTCCATTAACCTTATATTTATGACACTGAAT 156

OY 1078 tga-----aaatagtctctctctctctctctctctctctctctctct 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 TGAATTTTCATTAATTTTCACATGTCACAAATATVCTCTTTGATTCTTCCAAAC 96

OY 1115 attaaacgttaaaactactcttagt 1143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 ATTTAAAAATGTAAAAACCATCTTAGT 67

```

```

RESULT 7
US-60-245-228-200
Sequence 200, Application US/60245228
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
FILE REFERENCE: C1000878
CURRENT APPLICATION NUMBER: US/60/245, 228
NUMBER OF SEQ ID NOS: 630
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 200
LENGTH: 196093
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(196093)
OTHER INFORMATION: n = A,T,C or G
US-60-245-228-200

```

```

Query Match 3.8%; Score 87; DB 63; Length 196093;
Best Local Similarity 70.1%; Pred. No. 9.7e-12;
Matches 117; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
OY 881 ttcccaagcctgagccaggttggcaaacctcttcctgtaagaacagaggaacta 940

```

```

Db 170687 ttgtlaaaacctaccacaggaattggcaaaccttctctgtaagaagcagagtaata 170746
OY 941 tttaagctctgtgagcctatggtctcagtcacaaactactatctctcctctagca 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170747 tttaagctctgtgagcctatggtctcagtcacaaactactatctctcctctagca 170806
OY 1001 cgaagcaattagcaacatatgtcacacaacatatgtgacccatg 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170807 tgaagcagcctatgagcctatgacataatgagtgagtgagcagatg 170853

```

```

RESULT 8
US-60-258-272-80
Sequence 80, Application US/60258272
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
FILE REFERENCE: C10001041
CURRENT APPLICATION NUMBER: US/60/258, 272
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 80
LENGTH: 196093
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(196093)
OTHER INFORMATION: n = A,T,C or G
US-60-258-272-80

```

```

Query Match 3.8%; Score 87; DB 64; Length 196093;
Best Local Similarity 70.1%; Pred. No. 9.7e-12;
Matches 117; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

```

OY 881 ttcccaagcctgagccaggttggcaaacctcttcctgtaagaacagaggaacta 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170687 ttgtlaaaacctaccacaggaattggcaaaccttctctgtaagaagcagagtaata 170746
OY 941 tttaagctctgtgagcctatggtctcagtcacaaactactatctctcctctagca 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170747 tttaagctctgtgagcctatggtctcagtcacaaactactatctctcctctagca 170806
OY 1001 cgaagcaattagcaacatatgtcacacaacatatgtgacccatg 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170807 tgaagcagcctatgagcctatgacataatgagtgagtgagcagatg 170853

```

```

RESULT 9
PCT-US01-15674A-420/C
Sequence 420, Application PC/TUS0115674A
GENERAL INFORMATION:
APPLICANT: JOAN, KNOEL H
APPLICANT: ROGAN, PETER K
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT
FILE REFERENCE: 30307
CURRENT APPLICATION NUMBER: PCT/US01/15674A
NUMBER OF SEQ ID NOS: 613
SOFTWARE: PatentIn version 3.1
SEQ ID NO 420
LENGTH: 224
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: repeat_region
LOCATION: (1)...(224)
OTHER INFORMATION: cheshire_a

```



Best Local Similarity 63.2%; Pred. No. 1.1e-10;  
Matches 144; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

```
OY 879 cgttccatgctgagcaggttgggcaaaccttctcgttaaaagaacagaggaac 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5023 cttagaagaagcttagagcaggggtcaacaaccttcttgaagaagaagtaaa 5082
OY 939 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 998
    ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5083 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 5142
OY 999 cagcaagaacattgacacacatagtcacgaacactactactactactact 1058
    ||||| ||||| ||||| ||||| ||||| ||| ||| ||| ||| ||| |||||
Db 5143 caccagaagcagctacagacata-ctcaagaagaagagcgtgtcacaataaatttact 5201
OY 1059 tattatgatacggaaacctgaaataatgtcttcttcttcttcttctt 1106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5202 tatgactgaattgaattcactatatacttcttcttcttcttcttctt 5249
```

## RESULT 13

```
US-60-233-468-1401
; Sequence 1401, Application US/60233468
; GENERAL INFORMATION:
; APPLICANT: Ring, Huijun Z.
; APPLICANT: Malsen, Gareth
; APPLICANT: Townley, David
; APPLICANT: Morris, MacDonald
; APPLICANT: Valdes, Ana
; TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes
; FILE REFERENCE: GX-0013-2 P
; CURRENT APPLICATION NUMBER: US/60/233,468
; CURRENT FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 2488
; SOFTWARE: PERL Program
; SEQ ID NO 1401
; LENGTH: 179040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GB:AC022732_3
US-60-233-468-1401
```

Query Match 3.6%; Score 83.2; DB 62; Length 179040;

Best Local Similarity 63.2%; Pred. No. 1.1e-10;  
Matches 144; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

```
OY 879 cgttccatgctgagcaggttgggcaaaccttctcgttaaaagaacagaggaac 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5023 cttagaagaagcttagagcaggggtcaacaaccttcttgaagaagaagtaaa 5082
OY 939 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 998
    ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5083 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 5142
OY 999 cagcaagaacattgacacacatagtcacgaacactactactactactact 1058
    ||||| ||||| ||||| ||||| ||||| ||| ||| ||| ||| ||| |||||
Db 5143 caccagaagcagctacagacata-ctcaagaagaagagcgtgtcacaataaatttact 5201
OY 1059 tattatgatacggaaacctgaaataatgtcttcttcttcttcttctt 1106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5202 tatgactgaattgaattcactatatacttcttcttcttcttcttctt 5249
```

## RESULT 14

```
US-60-313-371-1401
; Sequence 1401, Application US/60313371
; GENERAL INFORMATION:
; APPLICANT: Ring, Huijun Z.
; APPLICANT: Malsen, Gareth
; APPLICANT: Townley, David
```

APPLICANT: Morris, MacDonald  
TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes  
FILE REFERENCE: GX-0013-5 P  
CURRENT APPLICATION NUMBER: US/60/313,371  
CURRENT FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 2447  
SOFTWARE: PERL Program  
SEQ ID NO 1401  
LENGTH: 179040  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: GB:AC022732\_3  
US-60-313-371-1401

Query Match 3.6%; Score 83.2; DB 70; Length 179040;

Best Local Similarity 63.2%; Pred. No. 1.1e-10;  
Matches 144; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

```
OY 879 cgttccatgctgagcaggttgggcaaaccttctcgttaaaagaacagaggaac 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5023 cttagaagaagcttagagcaggggtcaacaaccttcttgaagaagaagtaaa 5082
OY 939 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 998
    ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5083 tatttaggcctgtgtgcacatagtcacgaacactactactctgcctctgtag 5142
OY 999 cagcaagaacattgacacacatagtcacgaacactactactactactact 1058
    ||||| ||||| ||||| ||||| ||||| ||| ||| ||| ||| ||| |||||
Db 5143 caccagaagcagctacagacata-ctcaagaagaagagcgtgtcacaataaatttact 5201
OY 1059 tattatgatacggaaacctgaaataatgtcttcttcttcttcttctt 1106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5202 tatgactgaattgaattcactatatacttcttcttcttcttcttctt 5249
```

## RESULT 15

```
US-60-212-664-256/c
; Sequence 256, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Spier, Gene
; APPLICANT: Greendberg, Simon
; APPLICANT: Radkin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
; FILE REFERENCE: C10000687
; CURRENT APPLICATION NUMBER: US/60/212,664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 256
; LENGTH: 163174
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(163174)
; OTHER INFORMATION: n - A,T,C or G
US-60-212-664-256
```

Query Match 3.6%; Score 83; DB 60; Length 163174;

Best Local Similarity 65.2%; Pred. No. 1.2e-10;  
Matches 122; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
OY 903 tgggcaactcttctctgtaaaagcagcaggaactattttaggcctgtgtgcacat 962
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92341 tggcgaactcttctctgtaaaagcagcaggaactattttaggcctgtgtgcacat 92282
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 05:14:32 ; Search time 97.23 Seconds  
(without alignments)  
5658.943 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatacgaactcaactatagg.....aaatccacaagccttgc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUTS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	71.8	3.2	7218	US-08-232-463-14	Sequence 14, Appl
2	67.6	3.0	6235	US-09-305-384-5	Sequence 5, Appl
3	67.6	3.0	6679	US-09-305-384-1	Sequence 1, Appl
4	62.2	2.8	178	US-08-991-789A-210	Sequence 210, App
5	62.2	2.8	178	US-09-062-451-210	Sequence 14, Appl
6	52.6	2.3	7218	US-08-232-463-14	Sequence 14, Appl
7	52	2.3	3786	US-08-975-762-42	Sequence 42, Appl
8	52	2.3	3786	US-09-293-028-42	Sequence 42, Appl
9	52	2.3	3786	US-09-106-562-42	Sequence 42, Appl
10	50.6	2.3	502	US-08-967-101-91	Sequence 91, Appl
11	50.6	2.3	502	US-08-592-541-91	Sequence 91, Appl
12	50.6	2.3	502	US-09-124-698-91	Sequence 91, Appl
13	50.6	2.3	502	US-09-127-480-91	Sequence 91, Appl
14	50.6	2.3	502	US-08-496-841C-91	Sequence 91, Appl
15	50.6	2.3	1883	US-08-967-101-10	Sequence 10, Appl
16	50.6	2.3	1883	US-08-967-101-153	Sequence 153, Appl
17	50.6	2.3	1883	US-08-592-541-10	Sequence 10, Appl
18	50.6	2.3	1883	US-08-592-541-153	Sequence 153, Appl
19	50.6	2.3	1883	US-08-888-077A-8	Sequence 8, Appl
20	50.6	2.3	1883	US-09-124-698-10	Sequence 10, Appl
21	50.6	2.3	1883	US-09-124-698-153	Sequence 153, Appl
22	50.6	2.3	1883	US-09-127-480-10	Sequence 10, Appl
23	50.6	2.3	1883	US-09-127-480-153	Sequence 153, Appl
24	50.6	2.3	1883	US-08-496-841C-10	Sequence 10, Appl
25	50.6	2.3	1883	US-08-496-841C-153	Sequence 153, Appl
26	48.2	2.2	80246	US-09-078-294-4	Sequence 4, Appl
27	48.2	2.2	80595	US-09-078-294-3	Sequence 3, Appl

## ALIGNMENTS

28	45.8	2.0	1734	2	US-08-858-052-2	Sequence 2, Appl
29	45.8	1.9	1734	2	US-09-200-284-2	Sequence 2, Appl
30	43.6	1.9	460	2	US-08-967-101-48	Sequence 48, Appl
31	43.6	1.9	460	2	US-08-592-541-48	Sequence 48, Appl
32	43.6	1.9	460	2	US-08-127-480-48	Sequence 48, Appl
33	43.6	1.9	460	4	US-08-496-841C-48	Sequence 48, Appl
34	42.8	1.9	59065	4	US-09-813-817-3	Sequence 3, Appl
35	42.8	1.9	289	4	US-09-007-005-17	Sequence 17, Appl
36	42.4	1.9	289	4	US-09-244-796-17	Sequence 17, Appl
37	42.4	1.8	703	4	US-09-313-300-6	Sequence 6, Appl
38	41.4	1.8	176373	3	US-09-128-155-17	Sequence 17, Appl
39	41.2	1.8	7505	4	US-09-078-294-13	Sequence 13, Appl
40	39.8	1.8	350	2	US-08-332-766A-20	Sequence 20, Appl
41	39.2	1.8	84495	4	US-09-797-906-3	Sequence 3, Appl
42	39.2	1.8	152331	3	US-09-128-155-16	Sequence 16, Appl
43	38.6	1.7	43795	3	US-08-742-185-101	Sequence 101, App
44	38.6	1.7	1732	1	US-08-480-346-2	Sequence 2, Appl
45	38.2	1.7	1732	1	US-08-480-346-2	Sequence 2, Appl

RESULT 1  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFLINER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOMPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)863-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pC-F1s  
US-08-232-463-14



```

? FILING DATE: 11-Dec-1997
? CLASSIFICATION: <unknown>
? ATTORNEY/AGENT INFORMATION:
?   NAME: Potter, Jane E. R.
?   REGISTRATION NUMBER: 33,332
?   REFERENCE/DOCKET NUMBER: 210121.419C3
? TELECOMMUNICATION INFORMATION:
?   TELEPHONE: (206) 622-4900
?   TELEFAX: (206) 682-6031
? INFORMATION FOR SEQ ID NO: 210:
?   SEQUENCE CHARACTERISTICS:
?     LENGTH: 178 base pairs
?     TYPE: nucleic acid
?     STRANDEDNESS: single
?     TOPOLOGY: linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 210:
?
? OS-08-991-789A-210

```

Query Match	2.8%	Score 62.2	DB 4	Length 178
Best Local Similarity	71.3%	Pred. No. 8.4e-10		
Matches	82	Conservative	0	Mismatches 33
				Indels 0
				Gaps 0
QY	891	cttgaccacaggttggcgaacatctctcgttaagaaacagacaggaactttagc	950	
Db	20	CTGAGCTAGAGGTTTGGCCAACTTTTTCATATAAGACCAAGAGATTAATTTCAGGCTT	79	
QY	951	tgtgtgcataatggtctcagltcaacaactaactatctctgtcctctgttagcagaa	1005	
Db	80	TGTGGTGTGCGACAGTCTCTTGCAACTACACGTACGCTGGCCATTGTGCACTTAGA	134	

```

RESULT 5
US-09-062-451-210
; Sequence 210, Application US/09062451
; Patent No. 6344550
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Avenue, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062.451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31.392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 210:
SEQUENCE CHARACTERISTICS:
LENGTH: 178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-062-451-210

```

[illegible]

RESULT 6  
 US-08-232-463-14/c  
 Sequence 14, Application US/08232463  
 Patent No. 5670367  
 GENERAL INFORMATION:  
 APPLICANT: DORNER, F.  
 APPLICANT: SCHEIFLINGER, F.  
 APPLICANT: FALKNER, F. G.  
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 1800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232.463  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/935.313  
 FILING DATE:  
 APPLICATION NUMBER: 26-AUG-1991  
 FILING DATE:  
 APPLICATION NUMBER: EP 91 114 300.6  
 FILING DATE: 26-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)835-9300  
 TELEFAX: (703)863-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: PT29PLT-FL5  
 US-08-232-463-14

Query March	2.3%	Score 52.6	DB 1:	Length 7218;
Best Local Similarity	9.0%	Pred. NO. 8	4e-06;	
Matches 28;	Conservative 162;	Mismatches 121;	Indels 0;	Gaps 0;
Qy	484	catcaetgectcttgacatcttatccatagtagttagtgaaggctcttgtaatgc	543	
Db	1490	CATCACTGTTATTACCTACTTATGCAAGTGTGTAAGAAAGAAATTTGTACRRR	1431	

OY 544 ctaatccacatggtggagaggggagtggtggagagtgctgtggtgccc 603  
Db 1430 RR 1371  
OY 604 taactctgagtgtaagactcggccctccaggaacaaagattcagctgtgagcgt 663  
Db 1370 RR 1311  
OY 664 atagccagagagctgtgcccagagtgcaagagatttctgtcttaagaaat 723  
Db 1310 RR 1251  
OY 724 aaacaacactgagtagagtagagaggggtgtgtgcccagagagatgggaaag 783  
Db 1250 RR 1191  
OY 784 tctccagaggg 794  
Db 1190 RRRRRRRRRRR 1180

RESULT 7  
US-08-975-762-42  
Sequence 42, Application US/08975762  
Patent No. 6207169  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,762  
FILING DATE: 21-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-622-4900  
TELEFAX: 206-682-6031  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3786 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-975-762-42

Query Match 2.3%; Score 52; DB 4; Length 3786;  
Best Local Similarity 67.2%; Pred. No. 8.7e-06;  
Matches 90; Conservative 0; Mismatches 40; Indels 4; Gaps 1;

OY 914 ttcctgtaagaacacagaaagaaactatttagctctgtgtccatagtct---tca 969  
Db 3296 TTTCTGTACTGCTCCAGACAGTAACATTGTGACTTTGTATGTATATGTCTCTCT 3355  
OY 970 gtcaaacactcatctctgtcctctgttagcagaaagaacttagcaacaataatgtcaaca 1029

Db 3356 GTTCAATATATGAACCTTCATTGTAGACAGCAAGCGGCTGCAGACAAATATGTAAACA 3415  
OY 1030 aacataatgtgacc 1043  
Db 3416 GATGAGCATGACTC 3429  
RESULT 8  
US-09-295-028-42  
Sequence 42, Application US/09295028  
Patent No. 6277381  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: McNeill, Patricia D.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS  
AND TREATMENT OF EHRLICHIA INFECTION  
FILE REFERENCE: 210121.439c4  
CURRENT APPLICATION NUMBER: US/09/295,028  
CURRENT FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 42  
LENGTH: 3786  
TYPE: DNA  
ORGANISM: Ehrlichia sp.  
US-09-295-028-42

Query Match 2.3%; Score 52; DB 4; Length 3786;  
Best Local Similarity 67.2%; Pred. No. 8.7e-06;  
Matches 90; Conservative 0; Mismatches 40; Indels 4; Gaps 1;

OY 914 ttcctgtaagaacacagaaagaaactatttagctctgtgtccatagtct---tca 969  
Db 3296 ttcctgtaagaacacagaaagaaactatttagctctgtgtccatagtctcttct 3355  
OY 970 gtcaaacactcatctctgtcctctgttagcagaaagaacttagcaacaataatgtcaaca 1029  
Db 3356 gtccaactactgacctcttcattctgttagcagaaagcgctgcagacaataatgtcaaca 3415  
OY 1030 aacataatgtgacc 1043  
Db 3416 gatgagcatgactc 3429

RESULT 9  
US-09-106-582-42  
Sequence 42, Application US/09106582  
Patent No. 6306402  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,582  
FILING DATE: 29-JUN-1998

THERAP



```

? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Maki, David J.
? REGISTRATION NUMBER: 31,392
? REFERENCE/DOCKET NUMBER: 210121.439C2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 206-622-4900
? TELEFAX: 206-682-6031
? INFORMATION FOR SEQ ID NO: 42:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3786 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
?
?
? US-09-106-582-42

```

Query Match	2.3%	Score 52	DB 4:	Length 3786:
Best Local Similarity	67.2%	Pred No. 8.7e-06:		
Matches	90	Conservative	0	Mismatches 40
				Indels 4
				Gaps 1
Qy	914	ttctctgaagaagaccagagaaactatttttagctctgltgacatagttc----	tca	969
Db	3296	ttttctgtgactgcgacagacgtaaacatttttgactttgtatgttatgctctttt		3355
Qy	970	gtcacactcctcatctctctgctctgttagcagaaagcaattgacaacatatgtcaaca		1029
Db	3356	gtttgcacactcctgacactcttccttcattgtttgacgaaagggcgctgacagcaaatatgttaaca		3415
Qy	1030	aacatattgacc		1043
Db	3416	gattgacatgacctc		3429

RESULT 10  
US-08-967-101-91  
: Sequence 91, Application US/08967101  
: Patent No. 5840540  
: GENERAL INFORMATION:  
: APPLICANT: ST. GEORGE-HYSLOP, PETER H  
: APPLICANT: ROMMENS, JOHANNA M  
: APPLICANT: FRASER, PAUL E  
: TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
: TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
: NUMBER OF SEQUENCES: 183  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: TESTA, HURWITZ & TRIBEAULT  
: STREET: High Street Tower - 125 High Street  
: CITY: Boston  
: STATE: Massachusetts  
: COUNTRY: U.S.A.  
: ZIP: 02110  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/967,101  
: FILING DATE: 10-NOV-1997  
: CLASSIFICATION: 435  
: PRIORITY APPLICATION DATA:  
: APPLICATION NUMBER: 08/592,541  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Pletcher, Edmund R.  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (617) 248-7000  
: TELEFAX: (617) 248-7100  
: INFORMATION FOR SEQ. ID NO.: 91:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 502 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-967-101-91

```

Query Match	2.3%	Score 50.6;	DB 2;	Length 502;
Best Local Similarity	64.3%	Pred. No. 6.8e-06;		
Matches 92;	Conservative 0;	Mismatches 49;	Indels 2;	Gaps 1.

Qy	917	ctgtlaagagcccaagaagaaactcttttagcgcgcgtgtgcatactgcttcagtaaca	916
Db	264	CTGTAAAGGCCAGACTG--AATATTTTAAAGCTATGAGGCAATATGTCCTCAGGCGCA	321
Qy	977	ctctactatctgcgcctctgtgtacagcaagcaattagcaacatatgtcaacaacatat	1032
Db	322	ACACCTCAACTGTGTGACTGTGTGTGTGTAAGCAGCAGCAGACATGTATTACCAAGAGAG	381
Qy	1037	gttagcccaatgaacatttattt	1059
Db	382	GTGTGACTTTTCCATGAAATT	404

```

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1
```

Query Match	2.3%	Score 50.6	DB 2	Length 502
Best Local Similarity	64.3%	Pred. No. 6.8e-06		
Matches	92	Conservative	0	Mismatches 49; Indels 2; Gaps 1
QY	917	ctgttaagaaccgcagcagaactattttagtgcgtgtgtgccatattgctctcagtcacaa	976	
DB	264	CTGTAAAAAGCCACACAGC--AATATTATTAAAGCTCATAGCGTCATATATGCTCCAGGGCAA	321	

Qy 977 ctactctctctcctctgtagcagaagcaattagcacaatctgtcaacaacatat 1036  
Db 322 ACACCTCAACTGCTACTGCTAGTGAAGGACAGCAGACAAATGTTATTAAACGAAGAG 381  
Qy 1037 gtgaccccatgaaaacttattt 1059  
Db 382 GTGCTCACTTCCCAATGAAGATT 404

## RESULT 12

US-09-124-698-91  
Sequence 91, Application US/09124698  
Patent No. 6117978  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: FRASER, PAUL E  
APPLICANT: ROMMENS, JOHANNA M  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,698  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-124-698-91

Query Match 2.3%; Score 50.6; DB 3; Length 502;  
Best Local Similarity 64.3%; Pred. No. 6.8e-06;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

Qy 917 ctgtaagaacacagacgaactattttaggcctcgtgtgcacataatggtccagtcacaa 976  
Db 264 CTGTAAAAGCCAGACACG--AATATTTTAAGCTCTATGGCTCATATGCTCTCCAGGCGCA 321  
Qy 977 ctactctctcctcctctgtagcagaagcaattagcacaatctgtcaacaacatat 1036  
Db 322 ACACCTCAACTGCTACTGCTAGTGAAGGACAGCAGACAAATGTTATTAAACGAAGAG 381  
Qy 1037 gtgaccccatgaaaacttattt 1059  
Db 382 GTGCTCACTTCCCAATGAAGATT 404

## RESULT 13

US-09-127-480-91  
Sequence 91, Application US/09127480  
Patent No. 6194153

GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: FRASER, PAUL E  
APPLICANT: ROMMENS, JOHANNA M  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/127,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-127-480-91

Query Match 2.3%; Score 50.6; DB 4; Length 502;  
Best Local Similarity 64.3%; Pred. No. 6.8e-06;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

Qy 917 ctgtaagaacacagacgaactattttaggcctcgtgtgcacataatggtccagtcacaa 976  
Db 264 CTGTAAAAGCCAGACACG--AATATTTTAAGCTCTATGGCTCATATGCTCTCCAGGCGCA 321  
Qy 977 ctactctctcctcctctgtagcagaagcaattagcacaatctgtcaacaacatat 1036  
Db 322 ACACCTCAACTGCTACTGCTAGTGAAGGACAGCAGACAAATGTTATTAAACGAAGAG 381  
Qy 1037 gtgaccccatgaaaacttattt 1059  
Db 382 GTGCTCACTTCCCAATGAAGATT 404

## RESULT 14

US-08-496-841C-91  
Sequence 91, Application US/08496841C  
Patent No. 6210919  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: FRASER, PAUL E  
APPLICANT: ROMMENS, JOHANNA M  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 175  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby, PC  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/496,841C  
APPLICATION NUMBER: US/08/496,841C  
FILING DATE: 28-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul F. Pehlner, Ph.D.  
REGISTRATION NUMBER: 35,135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237

INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 91:  
US-08-496-841C-91

Query Match  
Best Local Similarity 2.3%; Score 50.6; DB 4; Length 502;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgttaagaacacagagaaactatttaggtctgtgtgcataatgtcagtcacaa 976  
DB 264 CTGTAAAGACGACACTG--AATATTTTAAAGCTCTATGGTCATATGGTCCAGGGCCAA 321  
QY 977 ctactcatctctgcctctgtgacgaagaacatlagcaacataatgtcacaacacat 1036  
DB 322 AACTCAACTGTGCTAGTGAAGGACGACGACACATGATTTAACAAGAGG 381  
QY 1037 gtgacccatgaaaactattt 1059  
DB 382 GTGTCACTTCCATGAAGTT 404

RESULT 15  
US-08-967-101-10/C  
Sequence 10, Application US/08967101  
Patent No. 5840540  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HORNITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/967,101  
FILING DATE: 10-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pletcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1883 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-967-101-10

Query Match  
Best Local Similarity 2.3%; Score 50.6; DB 2; Length 1883;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgttaagaacacagagaaactatttaggtctgtgtgcataatgtcagtcacaa 976  
DB 239 CTGTAAAGACGACACTG--AATATTTTAAAGCTCTATGGTCATATGGTCCAGGGCCAA 182  
QY 977 ctactcatctctgcctctgtgacgaagaacatlagcaacataatgtcacaacacat 1036  
DB 181 AACTCAACTGTGCTAGTGAAGGACGACGACACATGATTTAACAAGAGG 122  
QY 1037 gtgacccatgaaaactattt 1059  
DB 121 GTGTCACTTCCATGAAGTT 99

Search completed: August 16, 2002, 05:15:29  
Job time: 15898 sec

Mon Aug 19 06:46:40 2002

us-09-515-369b-1\_copy\_1\_2240.rni